

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/11, 15/00, 15/63, C07H 21/02, 21/04		A1	(11) International Publication Number: WO 00/04140 (43) International Publication Date: 27 January 2000 (27.01.00)
(21) International Application Number: PCT/US99/15849 (22) International Filing Date: 14 July 1999 (14.07.99) (30) Priority Data: 60/092,921 15 July 1998 (15.07.98) US 60/092,922 15 July 1998 (15.07.98) US 60/092,956 15 July 1998 (15.07.98) US (71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). KOMATSOULIS, George [US/US]; 9518 Garwood Street, Silver Spring, MD 20901 (US). DUAN, Roxanne, D. [US/US]; 5515 Northfield Road, Bethesda, MD 20817 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). MOORE, Paul, A. [US/US]; 19005 Leatherbark Drive, Germantown, MD 20874 (US). SHI, Yang-gu [CN/US]; Apartment 102, 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 3142 Quesada Street, N.W., Washington, DC 20015 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne		Terrace #316, Gaithersburg, MD 20878 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; Apartment 115, 410 Van Dyke Street, St. Paul, MN 55119-4321 (US). FLORENCE, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). MUCENSKI, Michael [US/US]; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). (74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
		Published With international search report.	
(54) Title: 71 HUMAN SECRETED PROTEINS			
(57) Abstract The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

71 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human

growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing

to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

In specific embodiments, the polynucleotides of the invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, or 7.5 kb in length. In a further embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of the coding sequence, but do not comprise all or a portion of any intron. In another embodiment, the nucleic acid comprising the coding sequence does not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene in the genome).

10 As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the
15 secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard,
25 Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to
30 sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an

overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

- 5 Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or
10 temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent
15 hybridization can be done at higher salt concentrations (e.g. 5X SSC).

- Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and
20 commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

- Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a
25 complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

- The polynucleotide of the present invention can be composed of any
30 polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of

single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation,

gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

5 (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

10 "SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Polynucleotides and Polypeptides of the Invention

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 1

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

30 PFCSGFFPSLWIYLPFIFNVSDLWMGSLSGCALPFCLXVFFLTVSPSAVGLLXF
AGGPLQTLFAWVSPVEAAEQRLLPVLSSGSFVSEGTQCQMPARALLYEVSVG

PYWEIPPSQDTRRSQTYLRRQSDP (SEQ ID NO: 195) . Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in pancreas islet cell tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the pancreas, including cancer and diabetes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For
10 a number of disorders of the above tissues or cells, particularly of the pancreas, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endocrine, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the
15 standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in tumors of pancreatic islet cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment and intervention of such tumors, in addition to other endocrine or
20 gastrointestinal tumors where expression has been indicated. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
25 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
30 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1099 of SEQ ID NO:11, b is an integer of 15 to 1113, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

10 following amino acid sequence:

HEGSCRAPGFSAHKGRGCPSRMTLPSRALASLGVGWGMRLRNQVTVSCG
GSRWSSRVALGAFSWVCGVALVLQPSGGGLGLTSPSEGCWEGELALAVLRA
PGGSPS (SEQ ID NO: 196). Polynucleotides encoding these polypeptides are also provided.

15 This gene is expressed equally in in .

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic disorders, particularly leukemia. Similarly,

20 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hemolymphoid,

25 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 104 as residues: Gly-29 to Ser-35, Ser-63 to Cys-68. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in hemangiopericytoma, breast lymph node, and bone marrow indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 969 of SEQ ID NO:12, b is an

integer of 15 to 983, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

5 The translation product of this gene shares sequence homology with the *Drosophila melanogaster* slit protein, a secreted protein that contains both an EGF domain and Leucine Rich Repeat domains. It is thought to be important in the development of midline glia and commissural axon pathways (See e.g., Rothberg et al. *Genes Dev.* 4:2169-87 (1990); which is hereby incorporated by reference herein).

10 This gene is expressed primarily in human hippocampus.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, and developmental disorders. Similarly, polypeptides and
15 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neurological, cancerous, or wounded tissues) or
20 bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 The tissue distribution within human hippocampus combined with the
25 homology to the *Drosophila* slit protein, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and
30 elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease,

Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 959 of SEQ ID NO:13, b is an integer of 15 to 973, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by

the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

IPLTLPGIFLLIRLFWRLGQSICGPGKLVLPQFCCGCAVISGHCVPRGMPSSW
LPGCFVLLCLVAVGCQLREWGVGGVSAVGLLALPHLQVLGMRGRGLISGG

5 (SEQ ID NO: 197) . Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 16. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 16.

This gene is expressed in KMH2 cells, osteoblasts, fetal spleen, Jurkat
10 membrane bound polysomes, breast, and cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer, immune, and skeletal disorders. Similarly, polypeptides and
15 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, skeletal, cancerous, or wounded tissues) or bodily
20 fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in KMH2 cells, osteoblasts, and fetal spleen indicates
25 that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in fetal spleen and T-cells indicates a role in the regulation of the
30 proliferation: survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved

in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene
5 product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory
10 bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity
15 disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have
20 commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify
25 agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
available and accessible through sequence databases. Some of these sequences are
25 related to SEQ ID NO: 14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
cumbersome. Accordingly, preferably excluded from the present invention are one or
30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1444 of SEQ ID NO: 14, b is an

integer of 15 to 1458, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

- 5 The translation product of this gene shares sequence homology with phospholipase A2 which cleaves fatty acids from carbon 2 of glycerol (ref. Prosite pattern documentation for PS2_HIS). Many snake venoms contain phospholipase A2, which prevents transmission of nerve impulses to muscles by blocking the release of acetylcholine from the neuron. Therefore, included in this invention as preferred
- 10 domains are Phospholipase A2 histidine active site domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). Phospholipase A2 is an enzyme which releases fatty acids from the second carbon group of glycerol. Structurally, PA2's are small and rigid proteins of 120 amino-acid residues that have four to seven disulfide bonds. PA2 binds a calcium ion which is required for activity.
- 15 The side chains of two conserved residues, a histidine and an aspartic acid, participate in a 'catalytic network'. Two different signature patterns for PA2's were developed. The first is centered on the active site histidine and contains three cysteines involved in disulfide bonds. The consensus pattern is as follows: C-C-x(2)-H-x(2)-C [H is the active site residue].
- 20 Preferred polypeptides of the invention comprise a Phospholipase A2 histidine active site domain selected from the following amino acid sequences: CCNQHDRC (SEQ ID NO: 199), SLTKCCNQHDRCYET (SEQ ID NO: 200) , and/or LTKCCNQHDRCYETCG (SEQ ID NO: 201) . Polynucleotides encoding these polypeptides are also provided. Further preferred are polypeptides comprising the
- 25 Phospholipase A2 histidine active site domain of the sequence listed in Table 1 for this gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is N-terminal or C- terminal to the Phospholipase A2 histidine active site domain. Alternatively, the additional contiguous amino acid residues is both N-terminal and
- 30 C-terminal to the Phospholipase A2 histidine active site domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The

above preferred polypeptide domain is characteristic of a signature specific to Phospholipase A2 proteins. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with Phospholipase A2 proteins. Such activities are known in the art, some of which are described elsewhere herein, or see, for example, McIntosh, et al. J. Biol. Chem. 270 (8), 3518-3526 (1995), incorporated herein by reference.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

GPAGKEAWIWSWLLPSPGPAPLPSASWGLCGDAPR
AAARGPVEPGAARMALLSRPALTLLLLLMAAVVRCQEQAQTTDWRATLTKTI
RNGVHKIDTYLNAALDLLGGEDGLCQYKCSDGSKPFPRYGYKPSPPNGCGSP
LFGXHLNIGIPSLTKCCNQHDRCYETCGKSKNDCDEEFQYCLSKICRDVQKTL
GLTQHVQACETTVELLFDSVIHLGCKPYLDSQRAACRCHYEKTDL (SEQ ID
NO: 198) . Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

This gene is expressed in a variety of cell types with no single type predominating.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological disorders, or metabolism disorders, specifically phospholipase A2 deficiencies. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neuromuscular system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., pancreas, cancerous and wounded tissues) or bodily fluids (e.g., bile, lymph,

serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 107 as residues: Gln-23 to Asp-30, Lys-66 to Cys-87. Polynucleotides encoding said polypeptides are also provided.
- The ubiquitous tissue distribution and homology to phospholipase A2 indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis
- 10 and treatment of neuromuscular disorders. Alternatively, considering the activity of phospholipase A2 as a block for neuro- transmission may suggest that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette
- 15 Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked
- 20 disorders, or disorders of the cardiovascular system. Alternatively, the homology to Phospholipase A2 proteins may indicate a potential use for the protein product of this gene in diagnosis, treatment and/or prevention of metabolism disorders, specifically deficiencies in Phospholipase A2. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate
- 25 ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly
- 30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
 5 formula of a-b, where a is any integer between 1 to 1991 of SEQ ID NO:15, b is an integer of 15 to 2005, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

10 In a specific embodiment polypeptides of the invention comprise the following amino acid sequence:

GTSSARPRGALPGGSAPSAPHGQLPGRAQPAPVSGPPPTSGLCHFDPAAPWPL
 WPGPWQLPPHPQDWPAPHDIPQDWVSFLRSFGQLTLCPRNGTVTGKWRGSH
 VVGLLTTLNFGDGPDRNKTRTFQATVLGSQMGLKGSSAGQLVLITARVTTER
 15 TAGTCLYFSAVPGILPSSQPPISCSEEGAGNATLSPRMGEECVSVWSHEGLVLT
 KLLTSEELALCGSRLLVLGSFLLFCGLLCCVTAMCFHPRRESHWSRTRL
 (SEQ ID NO: 202) . Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by
 20 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

ARAPPGEGLSPEAQPLLPMGNCQAGHNLHLCLAHHPPLVCATLILLGLS
 GLGLGSFLLTHRTGLRT LTSPRTGSLF (SEQ ID NO: 203) . Polynucleotides
 encoding these polypeptides are also provided.

25 This gene is expressed in a wide variety of tissue types including testes, cerebellum, dendritic cells, breast cancer, umbilical vein endothelial cells, epididymus, corpus colosum, chronic synovitis, liver hepatome, normal breast, osteoblasts, melanocytes, B cell lymphomas, and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
 30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, cancer, particularly of endothelial tissues. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system,

5 expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endothelial, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in

10 healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 108 as residues: Thr-52 to Gly-57. Polynucleotides encoding said polypeptides are also provided.

Expression within embryonic tissue and other cellular sources marked by

15 proliferating cells indicates that the protein product of this gene may play a role in the regulation of cellular division and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Additionally, the expression

20 in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic

25 development also relies on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain

30 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have

applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 929 of SEQ ID NO:16, b is an integer of 15 to 943, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

RFLSVXPQXEVFLLHPCVCFXGGHPSLLPDCRAVGGGWEAPRCCLHEALC
 QSLGCKAEEIVSVSESSAQRWCWYLLRGRKAGGRGPASPVLFALMRLESLCH
 LCLACLFFRLPATRTVYCMNEAEIVDVALGILIESRKQXKACEQPALAGADNP
 EHSPPCSVSPHTSSGSSSEEDSGKQALXPGLSPSRPGGSSSACSRSPEEEE

- 5 EEDVLKYVREIFFS (SEQ ID NO: 204) . Polynucleotides encoding these polypeptides are also provided. Polynucleotides of the invention do not consist of the nucleic acid sequences shown as GeneSeq Accession Nos: V59595 and V59744, which are hereby incorporated herein by reference.

- 10 This gene is expressed primarily in a variety of immune cell types, including stromal cells, dendritic cells, leukocytes, activated T-cells, macrophages, monocytes, neutrophils and to a lesser extent in a variety of other adult and fetal tissues.

- 15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain
- 20 tissues or cell types (e.g., immune, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 25 The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in fetal
- 30 tissue and various hematopoietic cancers indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all

hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

5 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, 10 such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. Moreover, the protein may represent a secreted 15 factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, 20 raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

 Many polynucleotide sequences, such as EST sequences, are publicly 25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or 30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1489 of SEQ ID NO:17, b is an

integer of 15 to 1503, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

- 5 When tested against Jurkat T-cell lines, supernatants removed from cells containing this gene activated the NF-kB (Nuclear Factor kB) pathway. Thus, it is likely that this gene activates T-cells through the NF-kB signal transduction pathway. NF-kB is a transcription factor activated by a wide variety of agents, leading to cell activation, differentiation, or apoptosis. Reporter constructs utilizing the NF-kB
- 10 promoter element are used to screen supernatants for such activity. Preferred polypeptides of the invention comprise the following amino acid sequence: VPGWPRACSPCQADSPRAHPPKLRGILRWAPVPLXCAALCPPLDSG MSMAACPEAPEPSFLREVSPSPASTQWHRPCNFRQVEANPRKEPKNLVWRD VSLGQXSRTPRGSGLELVRVCGGGMQRDKTVVEERVGEERERERERESLGG
- 15 AGKHGEMRCVYVRESVGAPGRAGGGGNGVNSVGCVRTVHSGSXPPPSAGV S (SEQ ID NO: 205). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in parts of the brain such as cerebellum and frontal lobe.

- Therefore, polynucleotides and polypeptides of the invention are useful as
- 20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of
- 25 the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene
- 30 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in cerebellum and frontal lobe indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, prevention and/or treatment of neurodegenerative disease states and behavioural disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1498 of SEQ ID NO:18, b is an

integer of 15 to 1512, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

- 5 In a specific embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:
- TRPGKELNLVFGQLSMARIGSTFVNMLMGWLYSKIEALLGSAGHTTLGITL
- 10 MIGGITCILSLICALALAYLDQRAERILHKEQGKTGEVIKLTVDKDFSLPLWLIF
IICVCYYVAVFPFIGLGKVFFTEKFGFSSQAASAINSVVYVISAPMSPVFGLLV
DKTGKNIWVLCA (SEQ ID NO: 206) . Polynucleotides encoding these polypeptides are also provided.

- The gene encoding the disclosed cDNA is believed to reside on chromosome
- 15 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in fetal tissue, and to a lesser extent in a variety of adult human tissues.

- Therefore, polynucleotides and polypeptides of the invention are useful as
- 20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fetal abnormalities, particularly developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For
- 25 a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, or cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having
- 30 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 111 as residues: Lys-30 to Thr-35. Polynucleotides encoding said polypeptides are also provided.

5 The tissue distribution in fetal tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

10 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Expression within embryonic tissue and other cellular sources marked by proliferating cells
15 indicates that this protein may play a role in the regulation of cellular division. Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating,
20 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in
25 proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders,
30 and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Furthermore, the protein

may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
5 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
10 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1641 of SEQ ID NO:19, b is an integer of 15 to 1655, where both a and b correspond to the positions of nucleotide
15 residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

The translation product of this gene shares sequence homology with human histiocyte-secreted factor (HSF) which is a novel cytokine that shows in vivo
20 antitumour activity without the cytotoxicity associated with tumour necrosis factor. Furthermore, The translation product of this gene also shares sequence homology with the human endogenous virus S71 gag polypeptide, the sequence of which is believed to represent a transformation locus for several cancers (See Genebank Accession No. [pir|A46312|A46312](#); all references available through this accession are
25 hereby incorporated by reference herein). Similarly, The translation product of this gene also shares homology with B219, a sequence that is expressed in at least four isoforms in very primitive hematopoietic cell populations which may represent a novel hemopoietin receptor (See, e.g., Cioffi, et al. Nat. Med. 2:585-589 (1996), which is hereby incorporated by reference herein). In a preferred embodiment
30 polypeptides of the invention comprise the following amino acid sequence:

CKDLCSR VYLLT LSP LLSYDPATSHSPRNTQ (SEQ ID NO: 207) . Also preferred are the polynucleotides encoding these polypeptides.

This gene is expressed primarily in tonsil, and colon, and to a lesser extent in a wide variety of human tissues.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, and gastrointestinal disorders, particularly tumors of the colon and tonsil. Similarly, polypeptides and antibodies directed to
10 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic, digestive and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, gastrointestinal, or
15 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 112 as residues: Met-1 to Cys-6. Polynucleotides encoding said polypeptides are also provided.

 The tissue distribution in tonsil and colon, combined with the homology to human histiocyte growth factor, the human endogenous viral protein, and B219
25 strongly indicate that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment and/or prevention, of a variety of hematopoietic and immune system disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19,
30 20, and 27, and elsewhere herein. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation

of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- 5 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2511 of SEQ ID NO:20, b is an

integer of 15 to 2525, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

5 The gene encoding the disclosed cDNA is believed to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

10 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

IICECWEEECQSCRLKITQPREICRMDFLVLFLFYLASVLMGLVLICVCSKTHS
LKGLARGGAQIFSCIEPCLQRAXHGLLHYLFHTRNHTFIVLHLVLQGMVYTE
YTWEVFGYCQELELSLHYLLPYLLLGVNLFFFTLTCGTNPGIITKANELLFLH
15 VYEFDEVMPKKNVRCSTCDLRKPARSKHCSVCNWCVHRFDHHCVVWNCCI
GAWNIRYFLIYVLTLTASAATVAIVSTTFLVHLVVMSDLYQETYIDDLGHLHV
MDTVFLIQYLFTFPRIVFMLGFVVVLSFLLGGYLLFVLYLAATNQTTNEWYR
GDWAWCQRCPLVAWPPSAEPQVHRNIHSHGLRSNLQEIFLPAFPCHERKKQE
(SEQ ID NO: 208) . Polynucleotides encoding these polypeptides are also provided.

20 This gene is expressed primarily in colon and brain and to some extent in all tissues.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and digestive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and digestive system, expression of this gene at significantly higher or lower levels is
30 routinely detected in certain tissues or cell types (e.g., neurological, gastrointestinal, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,

synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly,
10 the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive
15 compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation,
20 neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Alternatively, expression of this gene in colon may indicate a role in the detection, prevention and/or treatment of colon disorders such as colon cancer, Crohn's Disease, ulcers, and digestive tract disorders in general. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to
25 isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1382 of SEQ ID NO:21, b is an integer of 15 to 1396, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

- 10 When tested against Reh cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activation site) pathway. Thus, it is likely that this gene activates B-cells through the Jaks-STAT signal transduction pathway. GAS is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.
- 15 Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. This gene maps to chromosome 7, and therefore, is used as a marker in linkage analysis for chromosome 7.
- 20 This gene is expressed primarily in brain, and in the developing embryo. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, behavioral, immune, and developmental disorders.
- 25 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and developmental systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developing, immune, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue
- 30

or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
5 epitopes shown in SEQ ID NO: 114 as residues: Lys-60 to Asn-67. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions.
10 Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia,
15 trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the tissue distribution in developing
20 embryo indicates that the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the biological activity within B-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of
25 immune system disorders. Activation of genes within B-cells indicates a role for this protein in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of
30 cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions; in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1055 of SEQ ID NO:22, b is an integer of 15 to 1069, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

The gene encoding the disclosed cDNA is believed to reside on chromosome 6. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 6.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

LLSFKIRGLRTEDAGWAQSSGGLCVRGDAFWMPSSSSGLGSPSRPPSSFLCL

LLLLPPAALALLFFLDFFPPRAAVSPFLPDHCSARQPRVWRRETLNRSASGL

GCWARSTEQGA VGVATGTVLDI SLPASCLSLWPPGPSGGI (SEQ ID NO: 209)

. Polynucleotides encoding these polypeptides are also provided.

In a specific embodiment polypeptides of the invention comprise the following amino acid sequence:

5 QLGLCLTSASLPPASRCGHQAPLGASDLSAHHSAPGFSDSYFTMSCQSSLRA
EILQCPLVPSVSPPTHL PQGRANKSSRASLPLLPQTHWCLFPSARGWRRGIQSG
LPPGGSCTSPRSPPQTLHQHITLVNHNTSYWQSPST (SEQ ID NO: 210),
HQPPCLLPLAVATRPLWGHLTCLPIILHLVSVTLTSPCLANQAFQGQRSYNAL
WCPLFLLLPTSPKGEQTNHPEPACPCFPKLTGVFSLQHVVGAEEFSQVFLVD
10 PVPVLDHLLKLFTSTSHLLIIPHIGKAPAPDSLL EELSLSLATHCKVAVARFT
(SEQ ID NO: 211) . Also preferred are the polynucleotides encoding these

polypeptides. Polynucleotides of the invention do not consist of the nucleic acid sequence shown as GeneSeq Accession No. X04377, which is hereby incorporated herein by reference.

15 This gene is expressed primarily in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, behavioral and neurological disorders. Similarly, polypeptides and
20 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, or cancerous and wounded tissues) or bodily
25 fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
30 epitopes shown in SEQ ID NO: 115 as residues: Pro-2 to Gly-7, Ser-10 to Ser-16,

Pro-52 to Val-62, Arg-64 to Ser-73. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1644 of SEQ ID NO:23, b is an

integer of 15 to 1658, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

- 5 The translation product of this gene was shown to have homology to the lysosomal mannosidase alpha-B protein (See Genebank Accession No. P34098) which is thought to be important in protein metabolism. One embodiment of this gene comprises polypeptides of the following amino acid sequence:
- MAAEGSRFSSQSPGLVDRQGPKCDPSRLVSPWGRHGLRILQIGHHHGRDGGQH
 10 EATHHLLRVLRAPRVGKADEGAVDSDPSTPLQLKHEAAHAEDHAQQVHVVR
 RRVVQGRVTFARRGLVPQHFRPPWVRHIVSGHSESKARSRLFRCRNRSFRR
 AS (SEQ ID NO: 212) , and/or
 RLVSPWGRHGLRILQIGHHHGRDGGQHEATHHLL RVLRA (SEQ ID NO: 213) .
- 15 An additional embodiment is the polynucleotides encoding these polypeptides. This gene maps to chromosome 19, and therefore, is used as a marker in linkage analysis for chromosome 19.

 This gene is expressed primarily in brain, placenta, fetal liver, and to a lesser extent in most tissues.

- Therefore, polynucleotides and polypeptides of the invention are useful as
- 20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, reproductive, and hepatic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For
- 25 a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, hepatic, or cancerous and wounded tissues) or bodily fluids (e.g., bile, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a
- 30 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 116 as residues: Asn-34 to Lys-42, Leu-60 to Trp-70. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution predominantly in brain indicates a role in the
5 detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Alternatively, the tissue distribution in liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of
10 liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma.

15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
20 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1063 of SEQ ID NO:24, b is an integer of 15 to 1077, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene is expressed primarily in spinal cord, Merkel cells, and adipose tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, disorders of the nervous and immune systems, particularly those disorders relating to the CNS involving lipid metabolism disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For
5 a number of disorders of the above tissues or cells, particularly of the nervous and immune systems and adipose tissue, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, immune, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
10 individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spinal cord, Merkel cells and adipose tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the
15 treatment and/or diagnosis of diseases the nervous systems, such as spinal cord injury, neurodegenerative diseases, muscular dystrophy or obesity. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
20 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
25 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1191 of SEQ ID NO:25, b is an integer of 15 to 1205, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 16

The translation product of this gene shares sequence homology with the human uncoupling protein-2 which is thought to be important in energy metabolism, obesity, and the predisposition of hyperinsulinemia (See Genebank Accession No. gi|1857278). Recently, another group published on this gene, designating it brain mitochondrial carrier protein-1 (BCMP1) (J Biol Chem 1998 Dec 18;273(51):34611-5). One embodiment of this gene comprises polypeptides of the following amino acid sequence: PTDVLKIRMQAQ (SEQ ID NO: 214) , and/or TYEQLKR (SEQ ID NO: 215) . An additional embodiment is the polynucleotides encoding these polypeptides. This gene maps to the X chromosome, and therefore, is used as a marker in linkage analysis for the X chromosome.

This gene is expressed primarily in manic depression brain tissue, epileptic frontal cortex, human erythroleukemia cell line, T-helper cells, and to a lesser extent in endothelial and amygdala cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system or hematopoietic/immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system or hematopoietic/immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, hemolymphoid, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 118 as residues: Ser-34 to Thr-39, Gln-198 to Leu-205. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neural tissues combined with the homology to the human uncoupling protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and/or treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's

5 Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of

10 developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Alternatively, given the homology to uncoupling proteins, the gene and/or its

15 translation product may also be used in the diagnosis, treatment, and/or prevention of thermogenesis disorders such as obesity, cachexia, and hyperinsulinemia. Uncoupling proteins dissipate the proton gradient created from the oxidation of fuels by the electron transport chain, thus releasing stored energy as heat. Dysfunction of thermogenesis can induce disorders such as obesity and cachexia. It is thought that

20 obesity may result from decreased thermogenesis in humans. Alternatively, cachexia is a metabolic state in which energy expenditure exceeds food intake, for example in anorexia nervosa. Uncoupling proteins is useful for the treatment and/or prevention of diseases and/or disorders involved with aberrant metabolic and thermogenic pathways. The following method provides for the determination of respiration

25 uncoupling activity of the polypeptides of the present invention, including fragments and variants of the full length proteins.

Briefly, yeast are transfected with an expression vector expressing polypeptide of the present invention as previously described by Bouillaud et al., EMBO J., 13:1990 (1994) (incorporated by reference herein in its entirety). Rates of growth in

30 liquid medium of transformed yeast are measured in the presence of galactose, which induces expression, as described in International Publication No. WO 98/31396

(incorporated by reference herein in its entirety). Instantaneous generation times are compared between the polypeptide of the present invention and appropriate controls. An in vivo decrease of membrane potential associated with uncoupling of respiration is analyzed by flow cytometry of yeast labeled with the potential sensitive probe

5 DiOC6 (3) (3,3'-dihexyloxacarbocyanine iodine, Molecular Probes, Eugene, OR). The ability of a polypeptide of the present invention to influence mitochondrial activity and uncouple respiration is thus determined.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

10 related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

15 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1660 of SEQ ID NO:26, b is an integer of 15 to 1674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

20 The translation product of this gene shares sequence homology with 55 kD deglycosylated zona pellucida protein which is known to be important in egg fertilization (See Genebank Accession No.R39356). Preferred polypeptides of the invention comprise the following amino acid

sequence:

25 RPRPSASSLARSASLLPAAHGXGVGGAGGGSSXLRSRYQQQLQNEEESGEPEQ
AAGDAPPPYSSISAESAHXFDYKDESGFPKPPSYNVATTLPYDEAERTKAEA
TIPLVPGRDEDFVGRDDFDDADQLRIGNDGIF (SEQ ID NO: 216) ,
RYQQLQNEEESGEPEQAAGD (SEQ ID NO: 217) , and/or
PGRDEDFVGRDDFDDADQLRIG (SEQ ID NO: 218) . Polynucleotides encoding

30 these polypeptides are also provided.

Preferred polypeptide fragments of the invention comprise the following amino acid sequence: MLTFFMAFLFNWIGFFLSFCLTTSAAGRYG AISGFGLSLIKWILIVRFSTYFPGYFDGQY WLWWVFLVLGFLFLRGFINYAKVRKMPET FSNLPRTRVLFIY (SEQ ID NO: 219). Polynucleotides encoding these polypeptides are also provided.

Preferred polypeptide variants of the invention comprise the following amino acid sequence: MKKSLENLNRLQVMLLHLTA AFLQRAQHXYFDYKDESGFPKPPSYNVATTLP SYDEAERTKAEATIPLVPGRDEDFVGRDDFDDADQLRIGNDGIFMLTFFMAFLF 10 NWIGFFLSFCLTTSAAGRYGAISGFGLSLIKWILIVRFSTYFPGYFDGQYWLW WVFLVLGFLFLRGFINYAKVR KMPETFSNLPRTRVLFIY (SEQ ID NO: 220), MLLHLTA AFLQRAQFSTYFPGYFDGQYWLWWVFLVLGFLFLRGFINYAKV RKMPETFSN LPRTRVLFIY (SEQ ID NO: 221), MLTFFMAFLFNWIGFFLSFCLT TSAAGRYGAISGFGLSLIKWILIVRFSTYFPAFMNSLSRSKRTPAGSESR CRTQ 15 RNNHLL (SEQ ID NO: 222), and/or MKKSLENLNRLQVMLLHLTA AFLQRAHXIL TTRMSLGFQSPHLM (SEQ ID NO: 223) . Polynucleotides encoding these polypeptides are also provided.

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter 20 element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells and JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the 25 Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in adult kidney, colon adenocarcinoma, and fetal brain, and to a lesser extent, ubiquitously expression in many tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as 30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, disorders of kidney, colon cancers, and central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the

5 renal and neural systems, and cancers, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., renal, neural, urogenital, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression

10 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 119 as residues: Cys-15 to Gly-36. Polynucleotides encoding said polypeptides are also provided.

15 The tissue distribution adult kidney, colon adenocarcinoma, and fetal brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of kidney diseases, colon cancers, and disorders of the central nervous system. Additionally, the homology to the zona pellucida protein indicates that the gene product is used for male contraceptive development, and

20 infertility diagnosis etc. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

25 related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

30 formula of a-b, where a is any integer between 1 to 1951 of SEQ ID NO:27, b is an

integer of 15 to 1965, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

- 5 The translation product of this gene shares sequence homology with the chicken transferrin receptor in addition to a human prostate-specific protein homolog (See Genebank Accession Nos. *pir*|JH0570|JH0570 and *gi*|2565338 (AF026380), respectively). This gene also shares significant homology with both the murine and the rat hematopoietic lineage switch 2 proteins (See Genbank Accession Nos. 10 *g*3169729 and *g*3851632, respectively), which are induced during an erythroid to myeloid lineage switch.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MTVM~~DPKQ~~MNVAAA VWAVVS~~YVV~~ADME~~EML~~ PRS (SEQ ID NO: 224). Polynucleotides encoding these polypeptides are also provided.

- 15 This gene is expressed primarily in fetal tissues, such as liver/spleen and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pre-natal disorders, anomalies, deficiencies. Similarly, polypeptides 20 and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing fetus, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, cancerous and wounded tissues) or bodily fluids 25 (e.g., amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic 30 epitopes shown in SEQ ID NO: 120 as residues: Arg-31 to Lys-37, Lys-58 to Glu-65,

Asp-157 to Gly-168, Ile-219 to Gly-225, Ala-260 to Ser-268, Thr-276 to Glu-282.
Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of pre-natal disorders, anomalies and deficiencies. The homology to the hematopoietic lineage switch 2 proteins indicates that The translation product of this gene is useful for the detection and/or treatment of immune system disorders. In addition, the homology to the transferrin receptor indicates that the translation product of the present invention may have utility in the regulation of iron metabolism as well as the numerous genes under the stringent control of physiologic iron levels. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1849 of SEQ ID NO:28, b is an integer of 15 to 1863, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

PRVRSREPVAGAPGCGTAGPPAMATLWGGLRLGSLLSLCLALSLLLLAHC
QTPPSDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTHIYLSILGLLL
LYMVYLTLEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLAARSRRANV

LNKVEYAQQRWKLQVQEQRKSVFDRHVVLS (SEQ ID NO: 225).

Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 72 - 88 of the amino acid sequence referenced in
 5 Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 89 to 167 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

A preferred polypeptide variant of the invention comprise the following amino
 10 acid sequence:

MATLWGGLRLGSLLSCLALSVLLAHCQTPPRISMSDVNVSA LPIKKNS
 GHIYNKNISQKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIII
 YLSILGLLLLYMVYLT LVEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLA
 RSRSRANVLNKVEYGTAAL EASSPRAAKSLSLTGMLSSANWGIEFKVTRKKQ

15 ADNWKGTDWVLLGFILIPC (SEQ ID NO: 226). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain tissue, and to a lesser extent in other cell types and tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
 20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and neurodegenerative diseases of the brain and nervous system, such as depression, schizophrenia, Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, mania, dementia, paranoia, addictive behavior, sleep
 25 disorders, epilepsy, transmissible spongiform encephalopathy (TSE), Creutzfeldt-Jakob disease (CJD). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher
 30 or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developmental, or cancerous and wounded tissues) or bodily fluids (e.g., amniotic

fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 121 as residues: Gln-110 to Pro-120, Val-152 to Val-159. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis
10 of developmental, degenerative and behavioral conditions of the brain and nervous system. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of schizophrenia, Alzheimer's Disease, Parkinson's
15 Disease, Huntington's Disease, Tourette Syndrome, transmissible spongiform encephalopathy (TSE), Creutzfeldt-Jakob disease (CJD), mania, depression, dementia, paranoia, addictive behavior, obsessive-compulsive disorder and sleep disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to
20 identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
25 related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
30 formula of a-b, where a is any integer between 1 to 1612 of SEQ ID NO:29, b is an

integer of 15 to 1626, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

5 The translation product of this gene shares sequence homology with a recently published gene Dysferlin, which is thought to be a skeletal muscle gene that is mutated in Miyoshi myopathy and limb girdle muscular dystrophy (See Genbank Accession No. g3600028, and Nat. Genet. 20 (1), 31-36 (1998)).

 This gene is expressed primarily in fetal liver, fetal heart tissue, and T-cells.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunodeficiency, tumor necrosis, lymphomas, auto-immunities, cancer, inflammation, anemias (leukemia) and liver disorders, vascular disorders, and
15 cancers (e.g., hepatoblastoma, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver and
20 immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hepatic, developmental, vascular, or cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, bile, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard
25 gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies
30 (e.g., AIDS), immuno-suppressive conditions (transplantation) and hematopoietic disorders. In addition this gene product is applicable in conditions of general

microbial infection, inflammation or cancer. Expression in liver may suggest a role for this gene product in the treatment and detection of liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Alternatively, the tissue distribution in fetal heart tissue indicates that the protein product of this gene is useful for the diagnosis and treatment of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Additionally, the homology to the dysferlin gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for diseases related to degenerative myopathies that are characterized by the weakness and atrophy of muscles without neural degradation; such as Duchenne and Becker's muscular dystrophies. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 591 of SEQ ID NO:30, b is an integer of 15 to 605, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene is expressed primarily in haemopoietic cells and tumor cells, such as pancreatic tumor tissue, and to a lesser extent in bladder cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, haemopoietic disorders, diseases of the renal and pancreatic systems, and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemopoietic, pancreatic, and renal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., pancreas, renal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the renal, pancreatic and haemopoietic systems, and cancers thereof. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 917 of SEQ ID NO:31, b is an integer of 15 to 931, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

This gene is expressed primarily in liver tissue, cancer cells and fetal lung tissue, and to a lesser extent in dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic disorders, diseases of developing systems and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetus, metabolic systems and cancers, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, metabolic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 124 as residues: His-44 to Gly-49. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the fetus, metabolic systems and cancers. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1393 of SEQ ID NO:32, b is an integer of 15 to 1407, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

This gene is expressed primarily in central nervous system tissues and cancers, such as endometrial tumors, and to a lesser extent in other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the CNS and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and cancerous tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 125 as residues: Tyr-16 to Ser-22, Asp-209 to Leu-215. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of diseases of the central nervous system, as well as cancers of tissues where expression of this gene has been observed, such as in
5 endometrial tumors. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive
10 disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein
15 may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of
20 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1512 of SEQ ID NO:33, b is an
25 integer of 15 to 1526, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

The translation product of this gene shares sequence homology with low-
30 density lipoprotein receptor (See Genbank Accession No. >dbj|BAA24580.1), which is thought to be important in the pathogenesis of atherosclerosis and other disorders.

The translation product of this gene also shares sequence homology with a rat homolog of the human CD94 (See Genbank Accession No. gb|AAC10220.1).

This gene is expressed primarily in macrophages, eosinophils, neutrophil and other cells of the haemopoietic and immune system.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune and haemopoietic systems and diseases of the endothelial and vascular system. Similarly, polypeptides and antibodies directed to
10 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, haemopoietic and vascular system, expression of this gene at significantly higher or lower levels is routinely detected in
15 certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
20 epitopes shown in SEQ ID NO: 126 as residues: Lys-9 to Ala-17, Met-55 to Leu-61, Tyr-105 to Cys-127, Asp-132 to Lys-141, Ser-165 to Tyr-172, Pro-178 to Met-186, His-222 to Gln-227. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution and homology to LDL receptor and rat CD94 homolog indicates that polynucleotides and polypeptides corresponding to this gene are useful
25 for the treatment and/or diagnosis of disorders of the immune, haemopoietic and vascular systems. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of hematopoietic disorders. This gene product is primarily expressed in hematopoietic cells and tissues, suggesting that it plays a role in the survival, proliferation, and/or
30 differentiation of hematopoietic lineages. Expression of this gene product in eosinophils and macrophage also strongly indicates a role for this protein in immune

function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1723 of SEQ ID NO:34, b is an integer of 15 to 1737, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 25

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

MAAAGRLPSSWALFSPLLAGLALLGVGPVPARALHNVTAEFGAEAWGTLA
AFGDLNSDKQTDLFVLRERNDLIVFLADQNAPYFKPKVKVSFKNHSALITSVV
20 PGDYDGDSQMDVLLTYLPKNYAKSELGAVIFWGQNQTLDPNNMILTILNRTFQ
DEPLIMDFNGDLIPDIFGITNESNQPQILLGGNLSWHPALTTTSMRIPSHAFI
DLTEDFTADLFLTTLNATTSTFQFEIWENLDGNFSVSTILEKPQNMVVGQSA
FADFDGDGHMDHLLPGCEDKNCQKSTIYLVRSGMKQWVPVLQDFSNGKTL
WGFVPFVDEQQPTEIPIPTLHIGDYNMDGYPDALVILKNTSGSNQQAFLLENV
25 PCNNASCEEARRMFKVYWELTDLNQIKDAMVATFFDIYEDGILDIVVLSKGY
TKNDFAIHTLKNNFEADAYFVKVIVLSGLCS NDCPRR (SEQ ID NO: 227).

Polynucleotides encoding these polypeptides are also provided.

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells and tissue cell types, through the JAK-STAT

signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the
5 GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in infant brain and placental tissues, and to a lesser extent in several other tissues including cancers.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, brain disorders and diseases of developing systems and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
15 type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and fetal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken
20 from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 127 as residues: Leu-56 to Thr-62, Gln-80 to Pro-87,
25 Gly-106 to Gln-113, Pro-122 to Lys-127, Gln-138 to Asn-146. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neural tissues and developing tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the central nervous system, disorders of
30 developing systems, and cancers. The tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful

for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product is produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product is produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 2228 of SEQ ID NO:35, b is an integer of 15 to 2242, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 26

Preferred polypeptides of the invention comprise the following amino acid sequence:

MTKREDGGYTFTATPEDFPKKHKAPVIDIGIANTGKFIMTASSDTTVLIWSLK
 GQVLSTINTNQMNNTAAVSPCGRFVASCGFTPDKVWEVCFGKKGEFQEV
 10 VRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTXVEYKKKQDPY
 LLKTGRFEEAAGAXPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
 GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNEST
 RQRLQQQLTQAQETLKS LGALKK (SEQ ID NO: 228). Polynucleotides encoding
 such polypeptides are also provided.

15 The gene encoding the disclosed cDNA is thought to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates
 20 myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS
 25 element, can be used to indicate proteins involved in the proliferation and differentiation of cells. Recently, another group published this gene, naming it WS beta-transducin repeats protein (See Human Genetics 103 (5), 590-599 (1998); which is hereby incorporated herein by reference), in which it was suggested that the protein is involved in William's Disease.

30 The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 12 - 28 of the amino acid sequence referenced in

Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the
5 the following amino acid sequence:

VIRHEGSTNMELSQMSXLMGLSVLLGLLALMATAAVXRGWLRAGEERSGRP
ACQKANGFPPDKSSGSKKQKQYQRIRKEKPQQHNFTHRLAAALKSHSGNIS
CMDFSSNGKYLATCADDRIRIWKSTKDFLQREHRSMRANVELDHATLVRFSF
10 DCRAFIVWLANGDTLRVFKMTKREDDGYTFTATPEDFPKKHKAPVIDIGIAN
TGK
FIMTASSDTTVLIWSLKGQVLSTINTNQMNTHAAVSPCGRFVASCGFTPDVK
VWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWK
LWDTXVEYKKKQDPYLLKTGRFEEAAGAXPCRLALSPNAQVLALASGSSIHL
15 YNTRRGEKEECFERVHGEICIANLSFDITGRFLASCGDRAVRLFHNTPGHRAM
VEEMQGHKLRASNSTRQLQQQLTQAQETLKS LGALKK (SEQ ID NO: 229).
Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in testes, synovial sarcoma and fetal tissues, and to a lesser extent in several other tissues.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the reproductive and developing systems and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
25 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, testicular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum,
30 plasma, seminal fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in testes tissue, synovial sarcoma, and fetal tissues, indicates that polynucleotides and polypeptides corresponding to this gene are useful
5 for the treatment and/or diagnosis of disorders of the reproductive and developing systems, and cancers. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment
10 of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues
15 of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.

Furthermore, the tissue distribution indicates that polynucleotides and
20 polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the
25 proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in
30 pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. This protein is useful for

the treatment, detection, and/or prevention of William's Disease. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
 5 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of
 10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2221 of SEQ ID NO:36, b is an
 15 integer of 15 to 2235, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

Preferred polypeptides of the invention comprise the following amino acid
 20 sequence positions 1-363, 2-363, 4-363, 5-363, 30-363, 31-363, 32-363, 75-363, 76-363 and 82-363 of the following amino acid sequence:
 MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTL
 FFAFECRYLAVQLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIE
 MEIEATNGAVPQGRPPRIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDN
 25 CVERFDHHCPWVGNCVGKRNYRYFYLFILSLSLLTIYVFAFNIVYVALKSLKI
 GFLETLKETPGTVLEVLCFFTLWSVVGLTGFHTFLVALNQTTNEDIKGSWTG
 KNRVQNPYSHGNIVKNCCEVLGGLPPSVLDRRGILPLEESGSRPPSTQETSSS
 LLPQSPAPTEL NSNEMPEDSSTPEEMPPPEPPEPPQEAEEAEK (SEQ ID NO:
 230). Polynucleotides encoding such polypeptides are also provided.

30 A preferred polypeptide variant of the invention comprises the following amino acid sequence: MLFLFSMATLLRTSFSDPGVIPRALPDEAA

FIEMEIEATNGAVPQGQRPPRIKNFQINNQIVKLKYCYTCKIFRPPRASHCSIC
 DNCVE RFDHHCPWVGNCVKGKRNRYFYLFILSLSLTIYVFAFNIVYVALK
 SLKIGFLETLKGNS WNC SRSPHLLLYTLVRRGTDWISYFPRGSQ PDNQ (SEQ
 ID NO: 231). Polynucleotides encoding these polypeptides are also provided.

5 This gene is expressed primarily in ovarian and endometrial tumors, fetal liver, spleen and brain tissues, and to a lesser extent in several other tissues and organs.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
 10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the developing systems, and cancers of the female reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above
 15 tissues or cells, particularly of the developing, female reproductive and fetal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a
 20 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 129 as residues: Pro-44 to Lys-54, Cys-88 to His-95, Val-103 to Tyr-108, Gln-181 to Ser-190, Thr-192 to Ile-206, Glu-233 to Ser-245, Ser-
 25 252 to Ala-286. Polynucleotides encoding said polypeptides are also provided.

 The tissue distribution in developing systems indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of developing and fetal systems, and cancers. Furthermore, the tissue distribution in ovarian and endometrial tumor tissues indicates that the
 30 translation product of this gene is useful for the detection, diagnosis, and/or treatment of cancers of the female reproductive system. Accordingly, preferred are antibodies

which specifically bind a portion of The translation product of this gene. Also provided is a kit for detecting these tumors. Such a kit comprises in one embodiment an antibody specific for The translation product of this gene bound to a solid support. Also provided is a method of detecting these tumors in an individual which
5 comprises a step of contacting an antibody specific for The translation product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. The above embodiments, as well as other treatments and diagnostic tests (kits and methods), are
10 more particularly described elsewhere herein.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
15 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2957 of SEQ ID NO:37, b is an integer of 15 to 2971, where both a and b correspond to the positions of nucleotide
20 residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

This gene is expressed primarily in normal and cancerous colon tissue, macrophages, endothelial cells and placental tissue, and to a lesser extent in several
25 other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, colon cancer and gastrointestinal disorders, immune disorders, vascular
30 diseases and disorders of developing systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, vascular and developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, gastrointestinal, developmental, 5 vascular, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 130 as residues: Thr-27 to Ser-33. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in macrophage, endothelial and placental tissues, and normal and cancerous colon tissues, indicates that polynucleotides and polypeptides 15 corresponding to this gene are useful for the treatment and/or diagnosis of immune, gastrointestinal and vascular disorders and diseases. Expression of this gene product in colon tissue indicates involvement in digestion, processing, and elimination of food, as well as a potential role for this gene as a diagnostic marker or causative agent in the development of colon cancer, and cancer in general. Accordingly, preferred are 20 antibodies which specifically bind a portion of the translation product of this gene. Also provided is a kit for detecting colon cancer. Such a kit comprises in one embodiment an antibody specific for The translation product of this gene bound to a solid support. Also provided is a method of detecting colon cancer in an individual which comprises a step of contacting an antibody specific for The translation 25 product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. The above embodiments, as well as other treatments and diagnostic tests (kits and methods), are more particularly described elsewhere herein. Alternatively, the tissue 30 distribution in placental tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders

of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product is produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta and endothelial cells also indicates that this gene product is produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Additionally, expression of this gene product in macrophage also strongly indicates a role for this protein in immune function and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1149 of SEQ ID NO:38, b is an integer of 15 to 1163, where both a and b correspond to the positions of nucleotide
5 residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

The translation product of this gene shares homology with HNK-sulfotransferase, which directs glycan synthesis (see Genbank Accession no.
10 AF033827).

This gene is expressed primarily in activated T cells, osteoclastoma, and glioblastoma, and to a lesser extent in various other normal and transformed cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation, immune defects, cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of
20 disorders of the above tissues or cells, particularly of the immune and hemopoietic systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative
25 to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 131 as residues: Pro-32 to Gly-48, Gln-63 to Thr-69, Pro-77 to Trp-84, Val-88 to Leu-94. Polynucleotides encoding said polypeptides are
30 also provided.

The tissue distribution in T-cells and various types of neoplasms indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, study and/or treatment of inflammatory and general immune defects, and various types of neoplasms. Expression of this gene product in T cells strongly indicates a role for this protein in immune function and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the tissue distribution in various cancerous tissues indicates that the translation product of the gene is useful for the detection, diagnosis, and/or treatment of these cancers, as well as cancers of other tissues where expression has been observed. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1918 of SEQ ID NO:39, b is an integer of 15 to 1932, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

Preferred polypeptides of the invention comprise the following amino acid sequence:

- 5 LHECLPGSISYLHPRTPLCLPPQHLSFSTFSPPWQPAMSPVPGTGGPPCGL
(SEQ ID NO: 232), and/or
MLPLLIICLLPAIEGKNCLRCWPELSALIDYDLQILWVTPGPPTELSQSIHSLFLE
DNNFLKPWYLDLDRDHLEETAKFFTQVHQAIKTLRDDKTVLLEEIYTHKNLFT
ERLNKISDGLKEKGAPPLHECLPGSISYLHPRTPLCLPPQHLSFSTFSPPWQP
10 AMSPVPGTGGPPCGL (SEQ ID NO: 233). Polynucleotides encoding these
polypeptides are also provided.

This gene is expressed primarily in infant brain, testes and activated T cells, and to a lesser extent in various other normal and transformed cell types.

- Therefore, polynucleotides and polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, neurological, reproductive and inflammatory conditions. Similarly,
polypeptides and antibodies directed to these polypeptides are useful in providing
immunological probes for differential identification of the tissue(s) or cell type(s). For
20 a number of disorders of the above tissues or cells, particularly of the neural, immune
and male reproductive systems, expression of this gene at significantly higher or
lower levels is routinely detected in certain tissues or cell types (e.g., neural, immune,
reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum,
plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken
25 from an individual having such a disorder, relative to the standard gene expression
level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic
epitopes shown in SEQ ID NO: 132 as residues: Gly-41 to Leu-46, Asp-67 to Thr-75,
30 Ile-114 to Ala-123. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution in infant brain tissue, testes tissue, and activated T-cells, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis, and/or treatment of neurological, reproductive and immune system disorders. Expression of this gene product in T-cells indicates a role
- 5 in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).
- 10 Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease,
- 15 sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell type. Alternatively, the tissue distribution in testes tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions
- 20 concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis
- 25 of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.
- 30 Furthermore, the tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the

detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 867 of SEQ ID NO:40, b is an integer of 15 to 881, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

The translation product of this gene shares sequence homology with some human and rodent melanoma and leukocyte specific antigens (see, for example, Genbank accession nos: gi|189384, gi|205898 and gi|180926). In addition, The translation product of this gene shares sequence homology with Tetraspan protein (see, for example, Genbank accession number: GI 3152703). Therefore, it is likely that the polypeptide of this gene shares some biological functions, such as cell-to-cell signaling, adhesion, proliferation, and differentiation with Tetraspan.

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 52-68 and 197 - 213 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed

that the protein product of this gene shares structural features to type IIIa membrane proteins.

The transmembrane 4 superfamily (TM4SF) or tetraspan superfamily has at least 16 members (including CD9, CD20, CD37, CD53, CD63, CD81, CD82, A15, 5 CO-029, Sm23, RDS, Uro B, Uro A, SAS, Rom-1, PETA3, and YKK8), is the second biggest subfamily among CD antigen superfamily. and activation antigen of T- cells. All TM4SF member contains four putative transmembrane domains, two extracellular loops, and two short cytoplasmic tails. They are variously expressed on Immature, early, mature, activated lymphocytes, monocytes, macrophages, granulocytes, 10 platelets, eosinophils, basophils, certain leukemic and lymphoma cells, and a variety of other cells and tissues. CD9 cell surface protein is expressed by both hematopoietic and neural cells, and may play a role for CD9 in intercellular signaling in the immune and nervous system. CD63 is a 53-Kd lysosomal membrane glycoprotein that has been identified as a platelet activation molecule, which play important role in cell 15 adhesion of platelets and endothelial cells. Increased mRNA for CD63 antigen was found in atherosclerotic lesions of Watanabe heritable hyperlipidemic rabbits, suggesting a potential role of CD63 in progression of atherosclerosis. CD63 is also a mast cell marker.

CD82 was originally identified as the target of several mAbs inhibitory to 20 syncytium formation induced by human T-cell leukemia virus type I (HTLV-I), the etiological agent of adult T-cell leukemia. Therefore, this gene could be a target for the development of a drug for this leukemia. CD81 is the target of an antiproliferative antibody. A diverse group of human cell lines, including hematolymphoid, neuroectodermal, and mesenchymal cells, express the CD81 protein. Many of the 25 lymphoid cell lines, in particular those derived from large cell lymphomas, were susceptible to the antiproliferative effects of the antibody. CD81 may therefore play an important role in the regulation of lymphoma cell growth. CD9, CD20, CD37, CD63, CD81 and CD82 have been implicated in the regulation of cell growth, adhesion, and signal transduction of B, T lymphocytes and some other non-lymphoid 30 cells. They associate with CD2, CD21, CD4, CD8, MHC Class II molecules, integrins, function as co-receptor for T, B and other lymphoid cells. Some

TM4SF are leukocyte antigens, highly expressed in activated leukocytes, lymphocytes, are highly specific surface marker for lymphoblastic leukemia, lymphoma, melanoma, and neuroblastoma. CD9 has been show to be involved in cell motility and tumor metastasis. These antigen could be a valuable immunogen or target to implement active and passive immunotherapy in patients with cancer. Others have been shown to be involved in inhibition of prostate cancer metastasis. This gene has close homology to C33 antigen (CD82), whic is a member of the transmembrane 4 superfamily (TMSF) and activation antigen of T- cells. C33 Ag (CD82 was originally identified as the target of several mAbs inhibitory to syncytium formation induced by human T-cell leukemia virus type I (HTLV-I), the etiological agent of adult T-cell leukemia. Therefore, this gene could be very important target for developing drug for leukemia. Other members of this family are Sm23, CO-029, R2, TAPA-1, CD9, CD37, CD53, and CD63. CD63 is a 53-Kd lysosomal membrane glycoprotein that has been identified as a platelet activation molecule.

There is strong evidence indicating that CD63 and Pltgp40, a platelet membrane glycoprotein are the same molecule and that CD63/Pltgp40 is identical to the well-characterized, stage-specific melanoma-associated antigen ME491. These antigen could be valuable immunogens or target to implement active and passive immunotherapy in patients with cancer.

This gene is expressed primarily in fetal tissue (kidney, heart, liver, spleen, brain), macrophages, dendritic cells, retina and to a lesser extent in various other tissues, mostly of lymphoid origin or epithelial cell types. In addition This gene is expressed in cancerous tissues (e.g. breast).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders and cancers in a variety of organs and cell types. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at

significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developmental, proliferating, immune, hematopoietic, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid, spinal fluid, or amniotic fluid) or another tissue or cell sample taken
5 from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 133 as residues: Tyr-123 to Tyr-131, Cys-134 to Ser-
10 145, Tyr-234 to Tyr-244. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution fetal cells and tissues and homology to tumor antigens indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, treatment and diagnosis of lymphoid and epithelial disorders and
15 neoplasms. Additionally, tissue distribution in immune cells and other tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting hematopoiesis, including cancers. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere
20 herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

25 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities,
30 such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity

disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits

5 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of

10 cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern

15 formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of

20 potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types

25 of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new

30 insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue

markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1918 of SEQ ID NO:41, b is an integer of 15 to 1932, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

The translation product of this gene shares limited sequence homology with VEGF which is thought to be important in regulation of endothelial cell growth. Therefore, it is likely that the protein encoded by this gene would share some similar
20 biological functions.

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells and tissue cell types, through the Jak-STAT signal transduction pathway. The gamma activating
25 sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and
30 differentiation of cells.

This gene is expressed in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, nervous system disease and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 134 as residues: Thr-25 to Pro-46. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to
5 identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
10 related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
15 formula of a-b, where a is any integer between 1 to 1150 of SEQ ID NO:42, b is an integer of 15 to 1164, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

20 The translation product of this gene shares sequence homology with human p150 which is thought to be important in signal transduction in neuronal cells. Therefore, it is likely that the protein encoded by this polynucleotide would share some similar biological functions with p150.

This gene is expressed primarily in whole embryo and cerebellum.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and growth defects/disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
30 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS, expression of this gene

at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of central nervous system, neurodevelopmental, cognitive, and memory disorders. The tissue distribution also indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked

by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1091 of SEQ ID NO:43, b is an integer of 15 to 1105, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

This gene is expressed primarily in PMA stimulated HL-60 cells and to a lesser extent in 6 week embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting cell differentiation, particularly hematopoietic disorders and/or defects. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 136 as residues: Pro-61 to Asp-68. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study of cellular differentiation and for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia. The tissue distribution also

indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the

5 expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Moreover, the protein may

10 represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Additionally, the expression within embryonic tissue and other

15 cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the

"Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere

20 herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain

25 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating,

30 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue

differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1248 of SEQ ID NO:44, b is an integer of 15 to 1262, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

This gene is expressed primarily in colon.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders and/or defects of the digestive tract including but not limited to cancers of the gastrointestinal tract. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at

significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., gastrointestinal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the digestive system particularly disorders involving the colon. Further, expression of this gene product in colon tissue indicates involvement in digestion, processing, and elimination of food, as well as a potential role for this gene as a diagnostic marker or causative agent in the development of colon cancer, and cancer in general. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the colon and/or other gastrointestinal tissue including, but not limited to, stomach, small intestine, large intestine, and rectum.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 503 of SEQ ID NO:45, b is an integer of 15 to 517, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

This gene is expressed primarily in blood cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, immune and hematopoietic diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 138 as residues: Pro-19 to Cys-29, Thr-35 to Glu-44, Val-72 to Lys-78. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis and/or treatment of disorders of the immune and hematopoietic system. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 844 of SEQ ID NO:46, b is an

integer of 15 to 858, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

5 This gene is expressed in multiple tissue systems such as brain, immune cells, prostate, uterus, testes, placenta, and fetal heart as well as in cancerous tissues such as ovarian tumors. .

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune, reproductive, urogenital, and central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
15 particularly of the central nervous system and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, reproductive, urogenital, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative
20 to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 139 as residues: Tyr-33 to Lys-38. Polynucleotides encoding said polypeptides are also provided.

25 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune, urogenital, reproductive, and central nervous systems. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis
30 of diseases of the central nervous system, as well as cancers of tissues where expression of this gene has been observed, such as in ovarian tumors. The tissue

distribution in central nervous system tissues also indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo. . Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in

proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution in uterus indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating female infertility. The protein product is likely involved in preparation of the endometrium of implantation and could be administered either topically or orally. Alternatively, this gene could be transfected in gene-replacement treatments into the cells of the endometrium and the protein products could be produced. Similarly, these treatments could be performed during artificial insemination for the purpose of increasing the likelihood of implantation and development of a healthy embryo. In both cases this gene or its gene product could be administered at later stages of pregnancy to promote healthy development of the endometrium. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution in testes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed

against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 6093 of SEQ ID NO:47, b is an integer of 15 to 6107, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

This gene is expressed in a wide range of tissue systems such as brain, immune cells, fetal liver, kidney, testes, breast, and pancreas as well as cancerous tissue such as ovarian tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system, immune system, urogenital, and reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, CNS, urogenital, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 140 as residues: Met-1 to Ser-7, Asp-32 to Pro-43, Ser-96 to Arg-102. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides
5 corresponding to this gene are useful for treatment and diagnosis of disorders of the immune, reproductive, urogenital and central nervous systems. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of diseases of the central nervous system, as well as cancers of tissues where
10 expression of this gene has been observed, such as in ovarian tumors. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome,
15 schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders.
20 Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation
25 of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such
30 as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities,

such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue

differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases.

The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 689 of SEQ ID NO:48, b is an integer of 15 to 703, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

This gene is expressed primarily in macrophages and fetal cells and to a lesser extent in cancerous ovarian tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune diseases, disorders of developing tissues, and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in

providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous
5 and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
10 corresponding to this gene are useful for treatment and diagnosis of developmental abnormalities and disorders of the immune systems. The tissue distribution cancerous ovaries indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of these tumors. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker
15 and/or immunotherapy target for the above listed tissues. Expression of this gene product in macrophage cells strongly indicates a role for this protein in immune function and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). This gene
20 product may have clinical utility in the treatment of immune dysfunction; in the correction of autoimmunity; in immune modulation; and in the control of inflammation.

The tissue distribution indicates polynucleotides and polypeptides
corresponding to this gene are useful for the diagnosis and treatment of a variety of
25 immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene
30 product is involved in the regulation of cytokine production, antigen presentation, or

other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the

polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of

5 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue

10 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution also indicates that polynucleotides and polypeptides corresponding to this gene are useful

15 for the treatment, diagnosis, and/or prevention of various skin disorders such as melanomas.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of

20 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 625 of SEQ ID NO:49, b is an

25 integer of 15 to 639, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

This gene is expressed primarily in neutrophils, bone marrow, brain, and fetal

30 cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic disorders, Limbic system disfunction/defects and disorders of the immune system and developing systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, Limbic system and developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 142 as residues: Ala-84 to Gln-93. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune, Limbic system, CNS and developing systems. Expression of this gene product in bone marrow, eosinophils, and neutrophils strongly indicates a role for this protein in hematopoiesis and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). This gene product may have clinical utility in the treatment of immune dysfunction; in the correction of autoimmunity; in immune modulation; and in the control of inflammation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune

Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury.

Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Additionally, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the

"Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 853 of SEQ ID NO:50, b is an

integer of 15 to 867, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

5 This gene is expressed primarily in ovary and to a lesser extent in fetal tissue, colon, and immune cells.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
10 not limited to, ovarian cancer, gastrointestinal and immune system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system, expression of this gene at significantly higher or lower
15 levels is routinely detected in certain tissues or cell types (e.g., reproductive, gastrointestinal, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
20 individual not having the disorder.

 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 143 as residues: Ile-23 to Ala-29. Polynucleotides encoding said polypeptides are also provided.

 The tissue distribution indicates that polynucleotides and polypeptides
25 corresponding to this gene are useful for diagnosis and treatment of ovarian cancer and related metastases. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating female infertility. The tissue distribution in colon tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders
30 involving the gastrointestinal tract. This may include diseases associated with digestion and food absorption, as well as hematopoietic disorders involving the

Peyer's patches of the small intestine, or other hematopoietic cells and tissues within the body. Similarly, expression of this gene product in colon tissue indicates again involvement in digestion, processing, and elimination of food, as well as a potential role for this gene as a diagnostic marker or causative agent in the development of

5 colon cancer, and cancer in general. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment,

10 and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

15 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have

20 applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue

25 differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the

30 protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their

interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
5 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
10 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1555 of SEQ ID NO:51, b is an integer of 15 to 1569, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a + 14.

15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 42**

The translation product of this gene shares sequence homology with retrovirus-related reverse transcriptase pseudogene. In addition, this gene shares homology with human interferon-beta (Genseq accession number T35524; all
20 references available through this accession are hereby incorporated herein by reference), therefore, it is likely that this gene and the protein encoded by this gene shares some similar biological functions with this protein.

This gene is expressed primarily in frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system,
30 expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily

fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution in frontal cortex and homology to retrovirus-related reverse transcriptase pseudogene and human interferon-beta indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurodegenerative diseases of the brain, particularly of the frontal cortex. The tissue distribution indicates polynucleotides and polypeptides
- 10 corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of
- 15 Alzheimer's Disease, Parkinson's Disease, multiple sclerosis, cystic fibrosis, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder,
- 20 learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

- Potentially, this gene product is involved in synapse formation,
- 25 neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show
- 30 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1182 of SEQ ID NO:52, b is an integer of 15 to 1196, where both a and b correspond to the positions of nucleotide
10 residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

This gene is expressed primarily in immune cells, brain, fetal tissue, and cancerous tissues (such as testes, stomach, lung, pancreas, ovaries) and to a lesser
15 extent in other numerous tissues including, but not limited to, testes and kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases. Similarly, polypeptides and antibodies
20 directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and immune cells expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids
25 (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
30 epitopes shown in SEQ ID NO: 145 as residues: Lys-23 to Lys-35, Met-46 to Tyr-52. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurodegenerative disorders of the frontal cortex, as well as, cancer or a number of tissues including but not limited to testes, stomach, lung, pancreas, and ovaries. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell

lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene
5 product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and
10 tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits
15 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their
20 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility
25 in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

30 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent

of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 931 of SEQ ID NO:53, b is an integer of 15 to 945, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 44

This gene is expressed primarily in epithelioid sarcoma and to a lesser extent in pancreatic carcinoma, aorta endothelial cells induced with TNF-alpha, and amniotic cells induced with TNF. This gene is also expressed, to a lesser extent, in cancerous lung and ovary tissue and fetal tissue.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, epithelioid sarcoma and related cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
10 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., amniotic, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another
15 tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 146 as residues: Tyr-39 to Arg-51. Polynucleotides
20 encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of certain cancers, including epithelioid sarcoma and pancreatic carcinoma. The tissue distribution in tumors of lung, ovary, and pancreas origins indicates that polynucleotides and
25 polypeptides corresponding to this gene are useful for the diagnosis and intervention of these tumors, in addition to other tumors where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked
30 by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of

developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including

blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene
5 product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory
10 bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity
15 disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits
20 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their
25 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
available and accessible through sequence databases. Some of these sequences are
25 related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 474 of SEQ ID NO:54, b is an

integer of 15 to 488, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

5 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:
PPVPPWISLPLTGSPPRPGFVPVSPFCFSPMTNGHQVLLLLLLTSAVAAGPWPO
10 VHAGQWGWMLPPGLPSVQARSGLGGLPGGPQWVPGGARGY (SEQ ID NO: 234). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in fetal and infant tissue, particularly infant brain and fetal liver/spleen libraries, and to a lesser extent in breast, ovary tumor, pharynx carcinoma, endometrial stromal cells, thymus, islet cell tumors, and adult
15 cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and other proliferative disorders. Similarly, polypeptides and
20 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and breast, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, developmental, hematopoietic, and cancerous and
25 wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The tissue distribution in developing cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis

and treatment of cancer and other proliferative disorders. The expression within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and
5 other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell
10 death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may
15 also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of
20 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue
25 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
 5 formula of a-b, where a is any integer between 1 to 2846 of SEQ ID NO:55, b is an integer of 15 to 2860, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

10 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

15 IQQWGDSVLGRRCDLLQLYLQRPELRVPVPEVLLHSEGAASSSVCKLDGLI
 HRFITLLADTSDSRALENRGADASMACRKLAVAHPLLLLRHLPMAALLHGR
 THLNQFEFRQQNHLSCFLHVLGLLELLQPHVFRSEHQGALWDCLLSFIRLLLN
 YRKSSRHAAAFINKFVQFIHKYITYNAPAAISFLQKHADPLHDLSFDNSDLVM
 LKSLLAGLSLPSRDDRTDRGLDEEGEEESSAGSLPLVSVSLFTPLTAAEMAPY
 MKRLSRGQTVEDLLEVLSDIDEMSRRRPEILSFFSTNLQRLMSSAECCRNLA
 20 FSLALRSMQNPSIAAAFLPTFMVCLGSQDFEVVQTALRNLPEYALLCQEHA
 AVLLHRAFLVGMYGQMDPSAQISEALRILHMEAVM (SEQ ID NO: 235).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in breast cancer, and to a lesser extent in a variety of other cancers, including uterine cancer, synovial sarcoma, and pharynx
 25 carcinoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast cancer; proliferative diseases and/or disorders. Similarly,
 30 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For

a number of disorders of the above tissues or cells, particularly of the breast, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, breast, proliferative, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, breast milk, urine, 5 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic 10 epitopes shown in SEQ ID NO: 148 as residues: Glu-35 to His-41, Ser-62 to Ala-67, Pro-145 to Leu-155, Glu-157 to Ser-163, Arg-190 to Val-197, Asp-208 to Pro-215, Ser-247 to Pro-252. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in breast cancer tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or 15 treatment of cancer. Elevated expression of this gene product in cancers, such as breast cancer, suggest that it is involved in the abnormal proliferation of cells, dedifferentiation, angiogenesis, and other processes that accompany the development of cancer. Thus, therapeutics targeted against this gene product is useful therapeutic products in and of themselves. Alternately, expression of this gene product at elevated 20 levels in breast tissue is reflective of expression within breast lymph nodes, and may suggest a hematopoietic role for this protein. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

25 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have 30 applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the

polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of

5 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue

10 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

15 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

20 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1545 of SEQ ID NO:56, b is an integer of 15 to 1559, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of this gene shares limited sequence homology with cytochrome-c oxidase. An alternative embodiment is the polypeptide comprising the following amino acid sequence:

MLLKHLQRMVSVPPQVKASALKVVTLTANDKTSVSFSSLPQGQVIYNVIVWD

30 PFLNTSAA YIPAHTYACSF EAGEGSCASLGRVSSKVFFTLFALLGFFICFFGHR

FWKTELFFIGFIIMGFFFYILITRLTPIKYDVNLILTA VTGSGVGMFLVAVVWR

FGILSICMLCVGLVLGFLISSVTFFTPLGNLKFHDDGVFWVTFSCIALIPVVF
 MGCLRILNLTGCVIGSYSVVLAIWSYWTSLSYITLNVLKRALNKDFHRAFTN
 VPFQTNDFIILAVWGMLAVSGITLQIRRERGRPFPPHPYKLWKQERERRVTNI
 LDPSYHIPPLRERLYGRLTQIKGLFQKEQPAGERTPLLL (SEQ ID NO: 236).

- 5 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

WARLRGPGAHARTSPQWRGPPSPAQAAMGFLQLLVXVLXSEHRVAGAAE

- 10 VFGNSSEGLIEFSVGKFRYF

ELNRPFPPEAILHDISSNVTFILFIQHSQYQNTTVSFSPRRRSPTM (SEQ ID NO: 237). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in keratinocytes, brain, and spinal cord and to a lesser extent in hematopoietic cells and tissues.

- 15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders; hematopoietic disorders; integumentary disorders; immune dysfunction; learning disabilities. Similarly, polypeptides and
 20 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., integumentary, neural, developmental,
 25 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 30 The tissue distribution in brain and spinal cord cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the

diagnosis and treatment of a variety of neurological and hematopoietic disorders. For example, elevated levels of expression of this gene product in brain and spinal cord indicates that it is involved in neurodegenerative disorders. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in

5 Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms,

10 hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

15 Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Alternately, expression of this gene product in hematopoietic cells indicates that it is involved in the proliferation, differentiation, survival, and activation of all hematopoietic lineages, including stem and progenitor cells. Expression of this gene

20 product in keratinocytes indicates that it is involved in normal skin function, and could be involved in skin disorders, dermatitis, and fibrosis. The protein is useful in detecting, treating, and/or preventing congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's Disease, basal cell carcinoma, squamous cell carcinoma,

25 malignant melanoma, Paget's Disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and

30 xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox,

molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athletes foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders (i.e., arthritis, trauma, tendonitis, chondromalacia and inflammation, etc.), autoimmune disorders (i.e., rheumatoid arthritis, lupus, scleroderma, dermatomyositis, etc.), dwarfism, spinal deformation, joint abnormalities, and chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2050 of SEQ ID NO:57, b is an integer of 15 to 2064, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 48

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

30 PRVRPASPPVRSPARWGSMAGSPLLWGPRAGGVGLLVLLLLGLFRPPPALCA
RPVKEPRGLSAASPPLARLALLAASGGQCPEVRRRGRCRPGAGAGASAGAER

QERARAEAQRLRISRRASWRSCCASGAPPATLIRLWAWTTTPTRLQRSSLALC
SAPALTLP (SEQ ID NO: 238). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in human pituitary and to a lesser extent in
5 pineal gland, and other areas of the brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pituitary dysfunction; abnormal growth; neurological defects;
10 insufficient milk secretion; abnormal smooth muscle contraction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and nervous systems, expression of this gene at significantly higher or lower levels is
15 routinely detected in certain tissues or cell types (e.g., endocrine, developmental, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, breast milk, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or
20 bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 150 as residues: Pro-36 to Gly-42, Pro-64 to Ala-76, Gly-83 to Ala-90, Ser-100 to Cys-108, Thr-126 to Ser-135. Polynucleotides encoding said polypeptides are also provided.

25 The tissue distribution primarily in pituitary cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of disorders. Elevated expression of this gene product in the pituitary indicates that it is possibly a hormone-like substance that either controls pituitary development itself, or various processes controlled by the
30 pituitary. These include growth, milk secretion, smooth muscle contraction, diuresis, blood pressure, and homeostasis. Thus, this gene product may have numerous clinical

applications. Expression of this gene product in other regions of the brain also indicates that it is involved in normal neurological function, and is useful in the treatment of a variety of neurological disorders. Representative uses are described in the "Biological Activity", "Hyperproliferative Disorders", and "Binding Activity" sections below, in Example 11, 17, 18, 19, 20 and 27, and elsewhere herein. Briefly, the protein can be used for the detection, treatment, and/or prevention of Addison's Disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothalamus, and testes. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1036 of SEQ ID NO:58, b is an integer of 15 to 1050, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

PRVRLATPNIWDL SMLFAFISLLV MLPTWWIVSSWL VVGVLVYLVIRALRL

PZ030PCT

WRTAKLQVTLKKYSVHLEDMATNSRAFTNLVRKALRLIQETEVISRGFTLVS
 AACPFNKAGQHPSQHLIGLRKAVYRTLRLANFQAARLATLYMLKNYPLNSES
 DNVNTNYICVVPFKELGLGLSEEQISEEEAHNFTDGFSLPALKVLFQLWVAQSS
 EFFRRLALLLSTANSPPGPLLTPALLPHRILSDVTQGLPHAHSACLEELKRSYE
 5 FYRYFETQHQSVQPCLSKTQQKSRELNNVHTAVRSLQLHLKALLNEVILEDE
 LEKLVCTKETQELVSEAYPILEQKLKLIQPHVQASNNCWEEAISQVDKLLRRN
 TDKKGKPEIACENPHCTVSTFEAAYSTHCRQRSNPRGAGIRSLCR (SEQ ID
 NO: 239). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane
 10 domain at about amino acid position 7 - 23 of the amino acid sequence referenced in
 Table I for this gene. Moreover, a cytoplasmic tail encompassing amino acids 24 to
 390 of this protein has also been determined. Based upon these characteristics, it is
 believed that the protein product of this gene shares structural features to type Ib
 membrane proteins.

15 The gene encoding the disclosed cDNA is believed to reside on chromosome
 12. Accordingly, polynucleotides related to this invention are useful as a marker in
 linkage analysis for chromosome 12.

This gene is expressed primarily in prostate and placenta and to a lesser extent
 in pancreatic tumors and hematopoietic cells.

20 Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, prostate cancer; pancreatic cancer; prostate dysfunction; hematopoietic
 disorders; reproductive diseases and/or disorders, and pancreatitis. Similarly,
 25 polypeptides and antibodies directed to these polypeptides are useful in providing
 immunological probes for differential identification of the tissue(s) or cell type(s). For
 a number of disorders of the above tissues or cells, particularly of the endocrine and
 immune systems, expression of this gene at significantly higher or lower levels is
 routinely detected in certain tissues or cell types (e.g., reproductive, prostate,
 30 pancease, placental, vascular, and cancerous and wounded tissues) or bodily fluids
 (e.g., lymph, serum, seminal fluid, plasma, urine, synovial fluid and spinal fluid) or

another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
5 epitopes shown in SEQ ID NO: 151 as residues: Pro-85 to Ser-94, Pro-127 to Thr-136, Glu-154 to Glu-160, Phe-240 to Ser-250, Leu-255 to Leu-265, Leu-341 to Lys-351, Thr-372 to Gly-384. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in prostate and placental cells and tissues indicates that
10 polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of reproductive disorders. Elevated expression of this gene product in the prostate indicates that it is involved in normal prostate function, and is a diagnostic marker for prostate cancer. Alternately, expression of this gene product in placenta indicates that it may play a role in normal vascular
15 function, and is involved in such processes as angiogenesis and endothelial cell chemotaxis. Thus, this gene product is useful in the treatment of myocardial infarction, cancer, ischemia, and diabetic retinopathy. Expression of this gene product in placenta may also be indicative of fetal health and development.

Similarly, expression of this gene product in hematopoietic cells indicates that
20 it is involved in the proliferation, differentiation, survival, or activation of all hematopoietic cell lineages. Finally, expression of this gene product in pancreatic cancers indicates that it may play a role in cancer in general, or in pancreatic function. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents
25 that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities. Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or
30 induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases

and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating
5 infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as,
10 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of
15 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2519 of SEQ ID NO:59, b is an
20 integer of 15 to 2533, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

When tested against Jurkat and K562 cell lines, supernatants removed from
25 cells containing this gene activated the GAS (gamma activating sequence) and ISRE (interferon-sensitive responsive element) promoter elements, respectively. Thus, it is likely that this gene activates myeloid, leukemia, and to a lesser extent, other immune or hematopoietic cells and tissue cell-types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes
30 which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.

Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. ISRE is also a promoter element found upstream in many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

AAPHPLLRPLCLWCPLWPAWPLRGRPRSAWKRWPLPVGPAKLGCSMTTR
QPTAVSWPCWLMSSSLSTACLAWTLTGSLAREATTRARSLPTWNC SARQV
PPSPPHSGLGRRGWAHCHLT CLLVTQLFRVGRIHPILSLPLVT (SEQ ID NO:
240). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, vascular diseases; aberrant angiogenesis; neurological disorders; learning disorders; placental insufficiency; and fetal distress. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and neurological systems (CNS/PNS), expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, reproductive, vascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 152 as residues: Met-1 to Thr-7, Glu-36 to Ser-43,
5 Pro-46 to Gly-63. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain and placental cells and tissues, combined with the detected GAS and ISRE biological activities, indicates that the protein products of this gene are useful for the diagnosis and/or treatment of a variety of neural,
10 reproductive, and vascular diseases and/or disorders. neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease,
15 Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism,
20 and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Expression of this gene product in placenta indicates that it may play a role
25 in blood vessel development or function, as the placenta is a highly vascularized organ. Thus, this gene product is involved in such processes as angiogenesis, endothelial cell chemotaxis, and vascular cord formation. Thus, it is useful in the treatment of such conditions as myocardial infarction; ischemia; and cancer. Alternately, expression of this gene product in the brain indicates that it may play a
30 role in the survival, proliferation, or function of neurons, and thus is useful in the diagnosis and treatment of such neurological disorders as ALS, schizophrenia, and

Alzheimer's Disease. It may likewise be involved in learning disorders as well. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement.

- 5 Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of
10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 885 of SEQ ID NO:60, b is an
15 integer of 15 to 899, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 51

In another embodiment, polypeptides comprising the amino acid sequence of
20 the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

LQLASQSAGIKGMSHCARPTFLTLLASCFWAAAIPNRNVILSVSFRPLHMQ
FTLSILVFILRILILRSFL (SEQ ID NO: 241). Polynucleotides encoding these

- 25 polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 40 - 56 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 57 to 60 of this protein has also been determined. Based upon these characteristics, it is
30 believed that the protein product of this gene shares structural features to type Ia membrane proteins.

This gene is expressed primarily in spleen derived from patients with chronic lymphocytic leukemia.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic lymphocytic leukemia; hematopoietic disorders; impaired immune function; cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spleen cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of hematopoietic disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Elevated expression of this protein in the spleens of patients with CLL indicates that it is a useful marker for this Disease. Alternately, it is associated with the development and/or progression of the disease, and is a useful target for therapeutic intervention. Additionally, this gene

product may play more general roles in hematopoiesis, and may serve to control cellular decisions regarding proliferation, survival, activation, and/or differentiation of all hematopoietic cell lineages. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate
5 ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
10 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
15 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1065 of SEQ ID NO:61, b is an integer of 15 to 1079, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:61, and where b is greater than or equal to a + 14.

20 **FEATURES OF PROTEIN ENCODED BY GENE NO: 52**

The translation product of this gene shares sequence homology with a putative tyrosine protein kinase from the Chilo iridescent virus. See, for example, Genbank accession no. gi|2738451 (AF003534). Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities
25 with tyrosine kinase and signaling proteins. Such activities are known in the art, some of which are described elsewhere herein.

This gene is expressed in a variety of tissues, including microvascular endothelial cells, dendritic cells, and fetal tissues. as well as several tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, vascular, immune, and developmental diseases and/or disorders, particularly cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above

5 tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., vascular, immune, developmental, proliferative, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having

10 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 154 as residues: Ala-21 to Lys-31, Arg-41 to Cys-56, Thr-92 to Cys-102, Arg-132 to Val-137, Lys-152 to Ile-159, Pro-199 to Ser-205, Arg-

15 210 to Asp-219, Ser-225 to Lys-230, Tyr-236 to Ala-241, Lys-243 to Leu-249, Thr-375 to Asp-381. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution and homology to a tyrosine kinase indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of cancer. Representative uses are described in the "Immune Activity"

20 and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes

25 suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such

30 as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities,

such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's

5 Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively,

10 protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons.

15 Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their

20 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

25 related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

30 formula of a-b, where a is any integer between 1 to 1914 of SEQ ID NO:62, b is an

integer of 15 to 1928, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

5 The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 2 - 18 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

 This gene is expressed primarily in neutrophils.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders, particularly cancer and immune suppression. Similarly, polypeptides and antibodies directed to
15 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids
20 (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred polypeptides of the present invention comprise immunogenic
25 epitopes shown in SEQ ID NO: 155 as residues: Gly-63 to Ser-72. Polynucleotides encoding said polypeptides are also provided.

 The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful as a marker for neutrophil monitoring in cancer and/or immune suppressed patients and/or during chemotherapy
30 or radiation therapy. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and

elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 767 of SEQ ID NO:63, b is an integer of 15 to 781, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

This gene is expressed primarily in IL-1 and LPS induced neutrophils, and to a lesser extent, in fetal brain.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, and neural diseases and/or disorders, particularly cancer and immune suppression. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
15 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and
20 spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 156 as residues: Ile-28 to Trp-37, Ser-68 to Lys-81.

25 Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful as a marker in neutrophils to monitor patients who are immune suppressed or cancer patients during chemotherapy or radiation therapy. Representative uses are described in the "Immune Activity" and
30 "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in

regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1180 of SEQ ID NO:64, b is an integer of 15 to 1194, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

This gene is expressed primarily in prostate.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, urogenital diseases and/or disorders, particularly prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urogenital system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., urogenital, prostate, renal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 157 as residues: Arg-30 to Gln-36. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in prostate cancer cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, treatment and diagnosis of prostate cancer and other urogenital disorders. Moreover, the expression within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis,

treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:65 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1663 of SEQ ID NO:65, b is an integer of 15 to 1677, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

A preferred polypeptide of the invention comprises the following amino acid sequence:

MVLVLRHPLCARERAFREPGRGLLTRTGQHDGAPAVTAVPGPLGAVAAAE
10 RRSAGWAGGSSPPRKVLWGDMRGRRAGVDVLGPALSSEAAGAEARGWGM
PGMGVGVGASETRGALFLGREGVHGPCPMDGLGPWPWGPW (SEQ ID NO:
242). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in rejected kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders affecting the kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of
20 disorders of the above tissues or cells, particularly of the urinary tract, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., urogenital, renal, kidney, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative
25 to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 158 as residues: Ala-30 to Gly-36, Asp-45 to Trp-50, Lys-65 to Cys-71, Pro-80 to Cys-87. Polynucleotides encoding said polypeptides are
30 also provided.

The tissue distribution in kidney indicates the protein product of this gene could be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. The protein is useful for modulating the immune response to aberrant proteins, as may exist in proliferating cells and tissues. Such modulation of the immune response would also show utility in inhibiting the rejection of transplanted tissues, particularly of the renal system. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1223 of SEQ ID NO:66, b is an integer of 15 to 1237, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:66, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

The translation product of this gene shares sequence homology with both human and mouse Fibulin-2 which is an extracellular matrix protein found in heart tissue (See Genbank Accession Nos. emb|CAA57876.1 and emb|CAA53040.1, respectively; all references available through these accessions are hereby incorporated

herein by reference; for example, J. Cell Biol. 123 (5), 1269-1277 (1993)). Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

MGPAVKMWTNAWKGLDDCHYNQLCENTPGGHRCSGPCRGYRMQGPSLPCL
DVNECLQLPKACAYQCHNLQGSYRCLCPPGQTLLRDGKACTSLERNQNVNT
5 TVSHRGPLLPWLRPWASIPGTSYHAWVSLRPGPMALSSVGRAWCPPGFIRQN
GVCTDLDECRVRNLCQHACRNTEGSYQCLCPAGYRLLPSGKNCQDINECEEE
SIECGPGQMCFNTRGSYQCVDTPCPATYRQGSPGTCFRRCSQDCGTGGPSTL
QYRLLPLPLGVRAHHDVARLTAFSEVGPANRTELSMLEPDPRSPFALRPLRA
GLGAVYTRRALTRAGLYRLTVRAAAPRHQS VFVLLIAVSPYPY (SEQ ID NO:
10 243). Polynucleotides encoding these polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

MRVLVVTIPIYWALARESGEALNGHSLTGGKFRQSHTWSLLQGAHDDPV
ARGLDPDGLLLLDVVVNGVVPGRWLTQIFKCR TLKKHYVQTRAWPAVRG
15 LHTALLPGRPPLVPTLQPQHPVQRGPGPPAPAGAAPAGLSYQLGL (SEQ ID
NO: 244). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the
20 following amino acid sequence:

HASGAFLVVRGEPQGSWGSMTGVINGRKFGVATLNTSVMQEAHSGVSSIHSS
IRHVPANVGPLMRVLVVTIPIYWALARESGEALNGHSLTGGKFRQESHVEF
ATGELLTMTQWPGVWIPMASCSTWWSMALSPDSLADADLQVQDFEEHYV
QTGPGQLFVGSTQRFFQGGLP SFLRCNHSIQYNAARGPQPQLVQHLRASAISS
25 AFDPEAEALRFQLATALQAEENEVGCPEGFELDSQGAFCVDVDECAWDAHL
CREGQRCVNLLGSYRCLPDCGPGFRVADGAGCEDVDECLEGLDDCHYNQLC
ENTPGGHRCSGPCRGYRMQGPSLPCLDVNECLQLPKACAYQCHNLQGSYRCL
CPPGQTLLRDGKACTSLERNQNVTTVSHRGPLLPWLRPWASIPGTSYHAWV
SLRPGPMALSSVGRAWCPPGFIRQNGVCTDLDECRVRNLCQHACRNTEGSY
30 QCLCPAGYRLLPSGKNCQDINECEEE SIECGPGQMCFNTRGSYQCVDTPCPAT
YRQGSPGTCFRRCSQDCGTGGPSTLQYRLLPLPLGVRAHHDVARLTAFSEV

GV PANRTELSMLEPDPRSPFALRPLRAGLGAVYTRRALTRAGLYRLTVRAAA
PRHQSVFVLLIAVSPYPY (SEQ ID NO: 245). Polynucleotides encoding these
polypeptides are also provided.

When tested against U937 and Jurkat cell lines, supernatants removed from
5 cells containing this gene repeatedly activated the GAS (gamma activating sequence)
promoter element. Thus, it is likely that this gene activates myeloid, T-cells, and to a
lesser extent, other immune and hematopoietic cells and tissue cell types, through the
JAK-STAT signal transduction pathway. GAS is a promoter element found upstream
of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway
10 is a large, signal transduction pathway involved in the differentiation and proliferation
of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of
the GAS element, can be used to indicate proteins involved in the proliferation and
differentiation of cells.

This gene is expressed primarily in kidney.

15 Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, diseases and/or disorders affecting the kidney and renal system.
Similarly, polypeptides and antibodies directed to these polypeptides are useful in
20 providing immunological probes for differential identification of the tissue(s) or cell
type(s). For a number of disorders of the above tissues or cells, particularly of the
urinary tract, expression of this gene at significantly higher or lower levels is
routinely detected in certain tissues or cell types (e.g., renal, urogenital, kidney, and
cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,
25 synovial fluid and spinal fluid) or another tissue or cell sample taken from an
individual having such a disorder, relative to the standard gene expression level, i.e.,
the expression level in healthy tissue or bodily fluid from an individual not having the
disorder.

Preferred polypeptides of the present invention comprise immunogenic
30 epitopes shown in SEQ ID NO: 159 as residues: Lys-32 to Ser-37, His-89 to Gly-94,
Asn-124 to Gln-130, Ala-163 to Val-168, Cys-196 to Arg-201, Gln-244 to Gln-264,

His-288 to Tyr-294, Leu-314 to Gln-319, Ala-392 to Ser-399, Pro-412 to Asp-419, Ala-452 to Pro-460, Arg-466 to Thr-473. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in rejected kidney, the homology to the conserved

5 Fibulin-2 protein, in addition to the detected GAS biological activity, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting kidneys, particularly proliferative disorders. Representative uses are described here and elsewhere herein. The protein product of this gene could be used in the treatment and/or detection of kidney diseases

10 including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Furthermore, the protein may also be used

15 to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

20 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

25 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1920 of SEQ ID NO:67, b is an integer of 15 to 1934, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

Preferred polypeptides of the invention comprise the following amino acid sequence:

MGEKFLLLAMKENHPECCKILKILHCMDPGEWLPQTEHCVHLPKEFLIWT
MDIASNERSEIQSVALLASKVISHHMQTCVENRELIAAELKQWVQLVILSCE
5 DHLTPTESRLAVVEVLTSTTFLTNPHPILELQDTLALWKCVLTLQSEEQAV
RDAATETVTTAMSQENTCQSTEFQVQDASIALALALAVLCDLLQQWDQL
APGLPILLGWLLGESDDLACVESMHQVEEDYLFEKAEVNFWAETLIFVKYL
CKHLFCLLSKSGWRPPSPPEMLCHLQRMVSEQCHLLSQFFRELPPAAEFVKTV
EFTRLRIQEERTLACRLAFLEGKEGEDTLVLSVWDSYAESRQLTLPRTEAA
10 C (SEQ ID NO: 246). Polynucleotides encoding such polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MGEPNRHPSM

FLLLLVLRLYASPMGTSSALSMGPFVPFIMRCGHSPVYHSREMAARALVP
FVMIDHIPNTIRTLSTL
15 PSCTDQCFRAPHSWGHFSRFFHLLQAYSDSKTRNEFRLPARAD (SEQ ID NO:
247). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

20 following amino acid sequence:

MTGREFFSRFPFLYLPFLKQLETVANTVDSMDGEPNRHPSMFLLLLVLRLY
ASPMGTSSALSMGPFVPFIMRCGHSPVYHSREMAARALVPFVMIDHIPNTIR
TLLSTLPSCTDQCFRQNHGHTLLQVFHLLQAYSDSKHGTNSDFQHELTDTV
CTKAKLWLAKRQNPCLVTRAVYDILFLLTCCLNRSKDNQPVLESIGFWEE
25 VRGIISGSELITGFPWAFKVPGLPQYLQSLRLAIAAVWAAAASGERETNVPI
SFSQLESAPFEVRSLTLEALLEKFLAAASGLGEKGVPLLNCNMGEKFLLLAM
KENHPECCKILKILHCMDPGEWLPQTEHCVHLPKEFLIWTMDIASNERSEIQ
SVALLASKVISHHMQTCVENRELIAAELKQWVQLVILSCEDHLTPTESRLAVV
EVLSTTFLTNPHPILELQDTLALWKCVLTLQSEEQAVRDAATETVTTAM
30 SQENTCQSTEFQVQDASIALALALAVLCDLLQQWDQLAPGLPILLGWLLG
ESDDLACVESMHQVEEDYLFEKAEVNFWAETLIFVKYLCKHLFCLLSKSG

WRPPSPEMLCHLQRMVSEQCHLLSQFFRELPPAAEFVKTVEFTRLRIQEERTL
 ACLRLLAFLEGKEGEDTLVLSVWDSYAESRQLTLPRTEAAC (SEQ ID NO:
 248). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have two transmembrane
 5 domains at about amino acid position 144 - 160, and 462 - 478 of the amino acid
 sequence referenced in Table 1 for this gene. Based upon these characteristics, it is
 believed that the protein product of this gene shares structural features to type IIIa
 membrane proteins. Included in this invention as a preferred domain is the formate
 and nitrite transporters domain, which was identified using the ProSite analysis tool
 10 (Swiss Institute of Bioinformatics). A number of bacterial and archaeobacterial
 proteins involved in transporting formate or nitrite have been shown [1] to be related:
 - *focA* and *focB*, from *Escherichia coli*, transporters involved in the bidirectional
 transport of formate. - *fdhC*, from *Methanobacterium formicicum* and
thermoformicicum, a probable formate transporter. - *nirC*, from *Escherichia coli* and
 15 *Salmonella typhimurium*, a probable nitrite transporter. - *Bacillus subtilis*
 hypothetical protein *yrhG*. - *Bacillus subtilis* hypothetical protein *ywcJ* (ipa-48R).
 These transporters are proteins of about 280 residues and seem to contain six
 transmembrane regions. As signature patterns, we selected two conserved regions.
 The first one is located in what seems to be a cytoplasmic loop between the second
 20 and third transmembrane domains; the second is part of the fourth transmembrane
 region. The 70 Kd yeast hypothetical protein YHL008c is highly similar, in its N-
 terminal section, to the prokaryotic members of this family. The consensus pattern is
 as follows: [LIVMA]-[LIVMY]-x-G-[GSTA]-[DES]-L-[FI]-[TN]-[GS].

Preferred polypeptides of the invention comprise the following amino acid
 25 sequence: IISGSELITG (SEQ ID NO: 249). Polynucleotides encoding these
 polypeptides are also provided. Further preferred are polypeptides comprising the
 formate and nitrite transporter domain of the sequence referenced in Table for this
 gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid
 residues of this referenced sequence. The additional contiguous amino acid residues is
 30 N-terminal or C- terminal to the formate and nitrite transporter domain. Alternatively,
 the additional contiguous amino acid residues is both N-terminal and C-terminal to

the formate and nitrite transporter domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to formate and nitrite transporter proteins. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with formate and nitrite transporter proteins. Such activities are known in the art, some of which are described elsewhere herein. It is believed that this gene maps to chromosome 2. Accordingly, polynucleotides derived from this gene are useful in linkage analysis as markers for chromosome 2.

10 This gene is expressed primarily in cells of the immune system, primarily T-cells and to a lesser extent in spleen, liver, thymus, tonsils, and testis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders, particularly disorders affecting hematopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of hematopoietic cells, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 160 as residues: Gly-2 to Pro-8, Ser-82 to His-92, Tyr-107 to Asp-117, Arg-162 to Pro-169, Ser-224 to Thr-229, Leu-310 to His-315, Ser-333 to Glu-338, Glu-381 to Ser-388, Gln-428 to Ala-433, Met-446 to Thr-455, Ser-548 to Ser-554, Gly-613 to Asp-618, Ser-627 to Gln-633. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting hematopoiesis, including cancers. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
 5 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3286 of SEQ ID NO:68, b is an integer of 15 to 3300, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where b is greater than or equal to a + 14.

10 **FEATURES OF PROTEIN ENCODED BY GENE NO: 59**

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

15 VDGIDKLDIEFLQQFLETHSRGPRHLHSPGHASQEATPGANMSSGTELLWPGAA
 LLVLLGVAASLCVRCSRPGAKRSEKIYQQRSLREDQQSFTGSRTYSLVGQAW
 PGPLADMAPTRKDKLLQFYPSLEDPASSRYQNFSKGSRHGSEEA YIDPIAMEY
 YNWGRFSKPPEDDDANSYENVLICKQKTTETGAQQEGIGGLCRGDLSSLAL
 20 KTGPTSGLCPSASPEEDEGI (SEQ ID NO: 250). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 10 - 26 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

25 The gene encoding the disclosed cDNA is believed to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

This gene is expressed primarily in bone marrow, CD34 positive cells, and immune cells, including, neutrophils, T-cells, B-cells, macrophages, monocytes, and
 30 dendritic cells and to a lesser extent in brain and tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting the immune and hematopoietic systems, particularly hematopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the the immune system and hematopoietic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 161 as residues: Ser-29 to Thr-57, Pro-74 to Lys-79, Pro-85 to Glu-107, Tyr-118 to Tyr-136, Gln-144 to Gln-152, Ala-182 to Glu-188. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune and hematopoietic cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting the immune system and hematopoiesis. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for

immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities.

Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors);

hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the
5 corresponding nucleic acid in gene therapy procedures.

Based upon the the proteins immune cell specific message distribution, it may be involved in many aspects of the immune response, especially its initial stages, inflammation, allograft rejection, infectious disease response etc. The expression of this clone is
10 frequently found in the hematopoietic cell cDNA libraries. Thus, this factor could be involved in the control of hematopoietic cell proliferation, differentiation, and function. Based on this one can postulate its use in the management of anemias, leukemias, neutropenia, thrombocytopenia, autoimmune diseases, blood tissue engraftment, and poikilothromerythromatosis. Furthermore, the
15 protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

20 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
25 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1783 of SEQ ID NO:69, b is an integer of 15 to 1797, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

5 VLWREASALVLSNRLSSGLLHDLQLQPAIHSRLFRRSRGLSEGEQSSVSLQRS
RVLSAMKHVNLNLYLLGVVLTLLSIFVRVMESLEGLLESPPGTSWTTRSQLAN
TEPTKGLPDHPSRSM (SEQ ID NO: 251). Polynucleotides encoding these
polypeptides are also provided.

This gene is expressed primarily in immune cells including activated T cells,
10 macrophages, jurkat cells, bone marrow cells, and osteoblasts and to a lesser extent in
kidney cortex, brain, placenta and lung.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, immune and hematopoietic diseases and/or disorders, particularly
inflammation and diseases related to inflammatory activity. Similarly, polypeptides
and antibodies directed to these polypeptides are useful in providing immunological
probes for differential identification of the tissue(s) or cell type(s). For a number of
disorders of the above tissues or cells, particularly of the immune system, expression
20 of this gene at significantly higher or lower levels is routinely detected in certain
tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded
tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal
fluid) or another tissue or cell sample taken from an individual having such a
disorder, relative to the standard gene expression level, i.e., the expression level in
25 healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
epitopes shown in SEQ ID NO: 162 as residues: Pro-34 to Met-63. Polynucleotides
encoding said polypeptides are also provided.

The tissue distribution in immune cells and tissues indicates that
30 polynucleotides and polypeptides corresponding to this gene are useful for treating or
diagnosing disease related to the normal or abnormal activation of T cells.

Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein.

Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1359 of SEQ ID NO:70, b is an integer of 15 to 1373, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

YTFHTQIFLDFPMIFLTVLPLAFLFLHSGFYHYISFSLFSLALFFFLDVATFR
RPGQLFCERSVLFDMFHFGFVSLFLHEWIQAKHFWAGLF
IVLPDVFVSVHLEAPDGSFPNIAKLSLIILLR (SEQ ID NO: 252).

Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 2 - 18 and 22 - 38 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIa membrane proteins.

This gene is expressed in many tissues including brain, liver, prostate, testes, cartilage, gall bladder. Expression is also seen in a number of tumors including colon carcinoma, pancreas tumor, osteoclastoma, ovarian cancer, B cell lymphoma and acute lymphocytic leukemias.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of various organs including the pancreas, colon, and bone. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the

major organs, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, hepatic, metabolic, reproductive, testicular, skeletal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or
5 another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in tumors and proliferative tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating or
10 diagnosing tumors of several major organs including the pancreas and large intestine. This protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below
15 and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain
20 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating,
25 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in
30 proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the

protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
5 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
10 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1565 of SEQ ID NO:71, b is an integer of 15 to 1579, where both a and b correspond to the positions of nucleotide
15 residues shown in SEQ ID NO:71, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

This gene is expressed primarily in dendritic cells and fetal liver/spleen and to a lesser extent in many tissues including tonsils, fetal lung, stromal cell lines, bone
20 marrow cell lines, placenta and tumors including hepatocellular carcinoma, pancreas tumor and osteosarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, diseases and/or disorders of the immune and hematopoietic system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is
30 routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,

synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution in dendritic cells and fetal liver/spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and treating disorders of the immune system particularly related to the control and generation of precursor cells. polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic
10 related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow
15 transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

- The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of
20 stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies
25 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of
30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1014 of SEQ ID NO:72, b is an integer of 15 to 1028, where both a and b correspond to the positions of nucleotide
5 residues shown in SEQ ID NO:72, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

This gene is expressed primarily in adrenal gland tumor and endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endocrine and vascular diseases and/or disorders, particularly diseases associated with the vascular endothelium. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
15 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endocrine, vascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or
20 cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in endothelial cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and treating
25 disorders that involve the vascular system including diseases such as atherosclerosis, neoangiogenesis associated with tumor growth and conditions associated with inflammation. Moreover, the protein is useful in the detection, treatment, and/or prevention of a variety of vascular disorders and conditions, which include, but are not limited to microvascular disease, vascular leak syndrome, aneurysm, stroke,
30 embolism, thrombosis, coronary artery disease, arteriosclerosis, and/or atherosclerosis. Alternatively, the protein is useful in the treatment, detection, and/or

prevention of metabolic disorders, particularly lethargy and depression. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3660 of SEQ ID NO:73, b is an integer of 15 to 3674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

The translation product of this gene is related to bovine PAM precursor. See Genbank record gi|163482 incorporated herein by reference. Moreover, see following patent publications are also incorporated herein by reference: J04311386 and WO8902460. Many bioactive peptides terminate with an amino acid alpha-amide at their COOH terminus. The enzyme responsible for this essential posttranslational modification is known as peptidyl-glycine alpha-amidating monooxygenase or PAM. An NH₂-terminal signal sequence and short propeptide precede the NH₂ terminus of purified PAM. The sequences of several PAM cyanogen bromide peptides were localized in the NH₂-terminal half of the predicted protein. The forms of PAM purified from bovine neurointermediate pituitary is generated by endoproteolytic cleavage at a subset of the 10 pairs of basic amino acids in the precursor. High levels of PAM mRNA have been found in bovine pituitary and cerebral cortex. In

corticotropic tumor cells, levels of PAM mRNA and pro-ACTH/endorphin mRNA are known to be regulated in parallel by glucocorticoids and CRF.

This gene is expressed primarily in endometrial tumors, dendritic cells, a multiple sclerosis library, kidney, hematopoietic cells, melanocytes, osteoblasts, the spleen, colon, ovary, stromal cells, fetal and adult brain, heart, and in tissues undergoing wound repair.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endometriosis, endometrial cancer, multiple sclerosis, hematopoietic diseases, bone disease, and wound healing. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly the hematopoietic system and female reproduction, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, immune, hematopoietic integumentary, skeletal, gastrointestinal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in dendritic and hematopoietic cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful as a therapeutic or diagnostic agent in diseases of hematopoietic origin as well as the female reproductive track due to the gene's primary pattern of expression. polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses

include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The protein may also have a very wide range of biological activities. Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 2783 of SEQ ID NO:74, b is an integer of 15 to 2797, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 65

The translation product of this gene shares sequence similarity with several G-protein coupled receptors (See Genbank Accession No. gb|AAC77910.1| (AF061443); all references available through this accession are hereby incorporated herein by reference; for example, Mol. Endocrinol. 12, 1830-1845 (1998)). G-protein coupled receptors are well known in the art and affect a variety of functions. In particular, the translation product of this gene shares similarity with Follicular Stimulating Hormone Receptor.

Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

15 GTRFPTGETPSLGFTVTLVLLNSLAFLLMAYIYTKLYCNLEKEDLSENSQSSMI
KHVAWLIFTNCIFFCPVAFFSFAPLITAISISPEIMKSVTLIFFP (SEQ ID NO:
253). Polynucleotides encoding such polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MIKHVAWLIFTNCIFFCP

20 VAFFSFAPLITAISISPEIMKSVTLIFFPCLLA (SEQ ID NO: 254). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

25 following amino acid sequence:

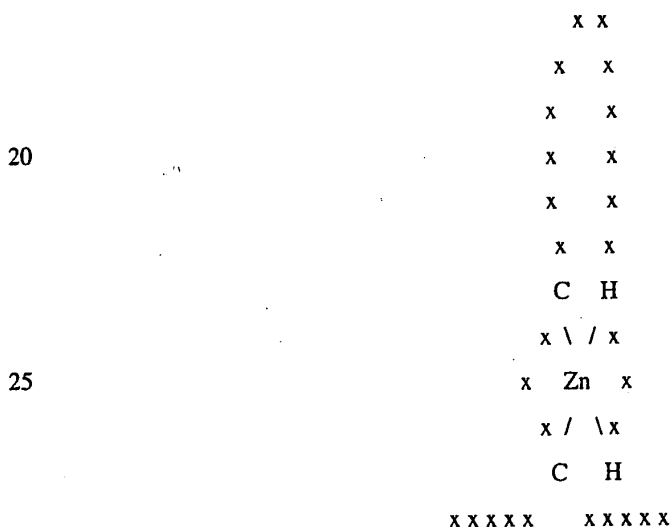
GTRFPTGETPSLGFTVTLVLLNSLAFLLMAYIYTKLYCNLEKEDLSENSQSSMI
KHVAWLIFTNCIFFCPVAFFSFAPLITAISISPEIMKSVTLIFFPLPACLNPVLYVF
FNPKFEDWKLLKRRVTKKSGSVSVSISSQGGCLEQDFYYDCGMYSHLQGN
LTVDCCESFLLTKPVSKHLIKSHSCPALAVASCQRPEGYWSDCGTQSAHS
30 DYADEEDSFVSDSDQVQACGRACFYQSRGFPLVRYAYNLPRVKD (SEQ ID
NO: 255). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 43 - 59 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 60 to 207 of this protein has also been determined. Based upon these characteristics, it is

5 believed that the protein product of this gene shares structural features to type Ia membrane proteins. Included in this invention as preferred domains are Zinc finger, C2H2 type domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). 'Zinc finger' domains [1-5] are nucleic acid-binding protein structures first identified in the *Xenopus* transcription factor TFIIIA. These

10 domains have since been found in numerous nucleic acid-binding proteins. A zinc finger domain is composed of 25 to 30 amino-acid residues. There are two cysteine or histidine residues at both extremities of the domain, which are involved in the tetrahedral coordination of a zinc atom. It has been proposed that such a domain interacts with about five nucleotides. A schematic representation of a zinc finger

15 domain is shown below:



30 Many classes of zinc fingers are characterized according to the number and positions of the histidine and cysteine residues involved in the zinc atom

coordination. In the first class to be characterized, called C2H2, the first pair of zinc coordinating residues are cysteines, while the second pair are histidines. A number of experimental reports have demonstrated the zinc- dependent DNA or RNA binding property of some members of this class. Some of the proteins known to include

5 C2H2-type zinc fingers are listed below. We have indicated, between brackets, the number of zinc finger regions found in each of these proteins; a '+' symbol indicates that only partial sequence data is available and that additional finger domains is present. In addition to the conserved zinc ligand residues it has been shown that a number of other positions are also important for the structural integrity of the C2H2

10 zinc fingers. The best conserved position is found four residues after the second cysteine; it is generally an aromatic or aliphatic residue. The consensus pattern is as follows: C-x(2,4)-C-x(3)-[LIVMFYWC]-x(8)-H-x(3,5)-H.

Preferred polypeptides of the invention comprise the following amino acid sequence: CDCCESFLLTKPVSCCKHLIKSH (SEQ ID NO: 256). Polynucleotides

15 encoding these polypeptides are also provided. Further preferred are polypeptides comprising the Zinc finger, C2H2 type domain of the sequence referenced in Table for this gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is N-terminal or C- terminal to the Zinc finger, C2H2 type domain.

20 Alternatively, the additional contiguous amino acid residues is both N-terminal and C-terminal to the Zinc finger, C2H2 type domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to zinc finger proteins. Based on the sequence similarity, the translation product of this gene is expected to

25 share at least some biological activities with G-coupled proteins, their receptors, and zinc finger proteins. Such activities are known in the art, some of which are described elsewhere herein.

This gene is expressed primarily in adult and fetal liver, human placenta, colon carcinoma cell lines and fibroblasts and to a lesser extent in the fetal and adult

30 brain, the developing nervous system, lung, pancreas, salivary gland, breast tissue, and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the liver, developmental abnormalities, neurologic diseases, lung cancer, pancreatic cancer, and colon cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological and hepatic origin, as well as the proliferation and/or differentiation of numerous types of tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hepatic, immune, hematopoietic, neural, gastrointestinal, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 167 as residues: Pro-62 to Asp-67, Arg-74 to Gly-80, Gln-146 to Glu-168. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in fetal liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for a diagnostic marker or therapeutic in a wide variety of disease states. polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the

5 differentiation and/or proliferation of various cell types. Alternatively, the protein expression in placental and brain tissue indicates the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble

10 factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. The protein is useful in the treatment, detection, and/or prevention of

15 bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's Disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic

20 depression, delirium, severe mental retardation and dyskinesias, such as Huntington's Disease or Gilles de la Tourette's syndrome. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed

25 against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:75 and may have been publicly available prior to conception of

30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2689 of SEQ ID NO:75, b is an integer of 15 to 2703, where both a and b correspond to the positions of nucleotide
 5 residues shown in SEQ ID NO:75, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by
 10 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

ALENSGSPGLQDSARAHFNXSLRSFSLRNQMYIFELSLYLEGTSFVVVLLFLL
 ISVSLDSPPTTKGWDSVLHIWVPLIVQ (SEQ ID NO: 257). Polynucleotides encoding these polypeptides are also provided.

15 This gene is expressed primarily in placenta and in hematopoietic cells, especially those of T-cell and monocyte origin and to a lesser extent in the brain, endothelial cells, and the lungs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
 20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic, vascular, and developmental diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the
 25 immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., vascular, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene
 30 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 168 as residues: Ser-30 to Trp-37. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in hematopoietic cells indicates that polynucleotides
5 and polypeptides corresponding to this gene are useful for therapeutic and/or diagnostic intervention in hematopoietic and developmental disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow
10 reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the
15 differentiation and/or proliferation of various cell types. Alternatively, the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble factor produced by smooth muscle that
20 regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue
25 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
5 formula of a-b, where a is any integer between 1 to 728 of SEQ ID NO:76, b is an integer of 15 to 742, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:76, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

10 This gene is expressed primarily in the prostate and to a lesser extent in human B-cell lymphomas.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, prostate cancer and diseases of hematopoietic origin, particularly of B-cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate and immune systems, expression of this gene at significantly higher or lower
20 levels is routinely detected in certain tissues or cell types (e.g., prostate, reproductive, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, seminal fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily
25 fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 169 as residues: Asp-33 to Lys-42. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in prostate tissue indicates that polynucleotides and
30 polypeptides corresponding to this gene are useful as a therapeutic or diagnostic marker for prostate cancer and disorders involving hematopoietic cells, especially

those of B-cell origin. Moreover, the expression within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. The protein is useful in modulating the immune response to aberrant proteins and polypeptides, as may exist in rapidly proliferating cells and tissues. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:77 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1811 of SEQ ID NO:77, b is an integer of 15 to 1825, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:77, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

GHESICGSCRSWIYFSIRCRRMRPWWSLLLEACATCAQTGPTRSTSCTQEVS
HSSSTAYPAPMRRRCCL PSPRSCT (SEQ ID NO: 258). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

This gene is expressed primarily in the brain and the developing embryo and to a lesser extent in the heart, colon, adipose tissue, kidney, mammary tissue, activated T-cells and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological diseases, developmental conditions, colon cancer, and hematopoietic diseases, especially of T-cell origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
10 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developmental, cardiovascular, adipose, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g.,
15 lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
20 epitopes shown in SEQ ID NO: 170 as residues: Thr-18 to Cys-26, Glu-29 to Thr-36, Ser-50 to Thr-55. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain, combined with the detected GAS biological activity, indicates that polynucleotides and polypeptides corresponding to this gene are useful for therapeutic and/or diagnostic agents in neurological diseases,
25 developmental abnormalities, colon cancer, and hematopoietic diseases, especially those of T-cell origin. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease,
30 Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal

cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated
5 expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity,
10 to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
15 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
20 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1660 of SEQ ID NO:78, b is an integer of 15 to 1674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 69

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 2 - 18 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein
30 product of this gene shares structural features to type II membrane proteins.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

- 5 K R A G V E V G G L V M A L A G S V F V L G G V L V L C V E R N G E G E M G W P Q H L P K S Q P L S
P P V A V R R C S F E R S W I D L L V E T S S S M V T C R Q Q V G T P N G M E G R G G G P K T T F P I R L
Q L S G A C A V R P E I Q W E V (SEQ ID NO: 259). Polynucleotides encoding these
polypeptides are also provided.

- 10 This gene is expressed primarily in activated monocytes, dendritic cells, and
in the tonsils.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders, particularly
15 leukemia, lymphomas, tumors of hematopoietic origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system, expression of this gene at significantly higher or lower levels is routinely detected in
20 certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 25 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 171 as residues: Gln-30 to Leu-38, Asn-75 to Thr-86. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution in activated monocytes, dendritic cells, and tonsils indicates that polynucleotides and polypeptides corresponding to this gene are useful
30 as a therapeutic and/or diagnostic agent for leukemias, lymphomas, and other diseases associated with cells of hematopoietic origin. Representative uses are described in the

"Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:79 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2177 of SEQ ID NO:79, b is an integer of 15 to 2191, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune cells and tissue cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 12.

This gene is expressed primarily in the placenta, brain, and liver and to a lesser extent in most other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic, neurological, vascular, and developmental diseases and/or disorders, particularly cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or

cell types (e.g., hematopoietic, neurological, vascular, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, bile, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful therapeutic and/or diagnostic agent in a multitude of disease states, particularly those involving the immune and neurologic systems. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the protein is useful in the detection, treatment, and/or prevention of a variety of vascular disorders and conditions, which include, but are not limited to microvascular disease, vascular leak syndrome, aneurysm, stroke, embolism, thrombosis, coronary artery disease, arteriosclerosis, and/or atherosclerosis. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement.

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:80 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1321 of SEQ ID NO:80, b is an integer of 15 to 1335, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 71

The translation product of this gene shares sequence homology with the murine Fig1 (interleukin-four induced gene 1) which shares homology to the monoamine oxidases, particularly in domains responsible for FAD binding. Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

QDWKAERSQDPFEKCMQDPDYEQLLKVTILEADNRIGGRIFTYRDQXTGWIG
 ELGAMRMPSSHRLHKLCQGLGLNLTKFTQYDKNTWTEVHEXKL RNYVVEK
 VPEKLG YALRPQEKGHSPEDIYQMALNQALKDLKALGCRKAMKKFERHTLL
 EYLLGEGNLSRPAVQLLGDMSEDGFFYLSFAEALRAXSCLSDRLQYSRIVG
 GWDLLPRALLSSLSGLVLLNAPVVAMTQGP HDVHVQIETSPARNLKV LKAD
 VLLTASGPAVKRITFS (SEQ ID NO: 260), and/or

LPRHMQEALRRLHYVPATKVFLSFRPFWREEHIEGGHSNTDRPSRMIFYPPP
 REGALLASYTWSDA AAFAGLSREEALRLALDDVAALHGPVVRQLWDGT
 GVVKRWAEDQHSQGGFVVQXPALWQTEKDDWTVPYGRIYFAGEHTAYPHG
 WVETAVKSALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVASSPSH
 DLAKEEGS (SEQ ID NO: 261). Polynucleotides encoding such polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

MAPLALHLLVLPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVTIL
EADNRIGGRIFTYRDQXTGWIGELGAMRMPSSHRLHKLCQGLGLNLTKFTQ
5 YDKNTWTEVHEXKLRNYVVEKVPEKLGALRPQEKGHSPEDIYQMALNQA
LKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSEDGFFYL
SFAEALRAXSCLSDRLQYSRIVGGWDLPLRALLSSLSGLVLLNAPVVAMTQG
PHDVH

VQIETSPARNLKVLKADVLLTASGPAVKRITFSPRCPATCRRRCGGCTTCR

- 10 PPRCS (SEQ ID NO: 262). Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with monoamine oxidases, disintegrins, metalloproteinases, and apoptosis modulating proteins. Such activities are known in the art, some of which are described elsewhere herein. Polynucleotides encoding these polypeptides are also provided.

- 15 The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 235 - 251 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 252 to 319 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia
20 membrane proteins.

This gene is expressed primarily in hematopoietic cells, particularly in dendritic cells, and activated monocytes and to a lesser extent in T-cells, endothelial cells, and cells associated with ulcerative colitis.

- Therefore, polynucleotides and polypeptides of the invention are useful as
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, leukemias, lymphomas, and diseases associated with antigen presenting cells, in addition to apoptosis dependant events. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
30 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression

of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 173 as residues: Gln-22 to Gln-44, Ala-90 to Gly-95, Lys-137 to Trp-146, Arg-171 to Asp-181, Glu-370 to Ser-380, Asp-447 to Gly-452, Gln-463 to Trp-469, Asn-504 to Ala-510, Asp-512 to His-519, Ala-541 to Val-550, Asn-558 to His-566. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution immune and hematopoietic cells and tissues, combined with the homology to the murine Fig 1 gene indicates that polynucleotides and polypeptides corresponding to this gene are useful as a therapeutic and/or diagnostic agent for hematopoietic diseases, especially those associated with antigen presenting cells. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's

Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1853 of SEQ ID NO:81, b is an integer of 15 to 1867, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:81, and where b is greater than or equal to a + 14.

Gene No.	cDNA Clone ID.	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HISCN02	209878 05/18/98	pSport1	11	1113	1	1113	232	232	103	1	26	27	106
2	HHGDM70	209878 05/18/98	Lambda ZAP II	12	983	102	983	69	69	104	1	57	58	86
3	HHPGO40	209878 05/18/98	Uni-ZAP XR	13	973	1	973	68	68	105	1	37	38	302
3	HHPGO40	209878 05/18/98	Uni-ZAP XR	82	984	1	984	74	74	174	1	37	38	224
4	HAMGG68	209878 05/18/98	pCMVSPORT 3.0	14	1458	1	1458	312	312	106	1	20	21	55
5	HAPOM49	209878 05/18/98	Uni-ZAP XR	15	2005	1	2005	251	251	107	1	22	23	189
5	HAPOM49	209878 05/18/98	Uni-ZAP XR	83	2664	1	2664	448	448	175	1	1	2	123
6	HBGBA69	209878 05/18/98	Uni-ZAP XR	16	943	1	933	62	62	108	1	38	39	60
7	HBJFJ26	209878 05/18/98	Uni-ZAP XR	17	1503	588	1480	290	290	109	1	26	27	128

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
7	HBJFJ26	209878 05/18/98	Uni-ZAP XR	84	1328	413	1305	591	591	176	1	20	21	59
8	HCEDH38	209878 05/18/98	Uni-ZAP XR	18	1512	1	1438	222	222	110	1	26	27	68
9	HDPOJ08	209878 05/18/98	pCMVSPORT 3.0	19	1655	1	1655	159	159	111	1	18	19	122
10	HDPRX82	209878 05/18/98	pCMVSPORT 3.0	20	2525	1	2525	128	128	112	1	32	33	82
11	HELK31	209878 05/18/98	Uni-ZAP XR	21	1396	25	1334	209	209	113	1	29	30	344
11	HCNUA40	97898 02/26/97 209044 05/15/97	pBluescript	85	1342	949	1237	960	960	177	1	33	34	105
12	HFPCX64	209878 05/18/98	Uni-ZAP XR	22	1069	1	1069	181	181	114	1	28	29	181
12	HFPCX64	209878 05/18/98	Uni-ZAP XR	86	1154	84	1154	257	257	178	1	28	29	87
12	HCEBW71	209225 08/28/97	Uni-ZAP XR	87	1197	141	1197	257	257	179	1	28	29	87

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
13	HFXDO60	209878 05/18/98	Lambda ZAP II	23	1658	1	1658	131	131	115	1	46	47	115
14	HHEPG41	209878 05/18/98	pCMVSPORT 3.0	24	1077	385	1043	514	514	116	1	35	36	70
14	HAUAI83	209626 02/12/98	Uni-ZAP XR	88	910	1	886	253	253	180	1	37	38	49
14	HJPAZ83	209626 02/12/98	Uni-ZAP XR	89	1076	398	1076		575	181	1	11	12	23
15	HKGHA42	209878 05/18/98	pSPORT1	25	1205	1	1205	143	143	117	1	21	22	63
16	HMIAP86	209878 05/18/98	Uni-ZAP XR	26	1674	13	1674	182	182	118	1	19	20	334
17	HMUAP70	209878 05/18/98	pCMVSPORT 3.0	27	1965	531	1914	183	183	119	1	16	17	221
17	HMUAP70	209878 05/18/98	pCMVSPORT 3.0	90	1842	407	1783	413	413	182	1	25	26	103
17	HAGFY16	97923 03/07/97 209071 05/22/97	Uni-ZAP XR	91	1963	209	1922	251	251	183	1	28	29	198

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
17	HBMCF37	209683 03/20/98	pBluescript	92	1487	79	1487	170	170	184	1	44	45	70
17	HFLQB16	209641 02/25/98	Uni-ZAP XR	93	1653	394	1637	413	413	185	1	25	26	82
17	HAGFY16	97923 03/07/97	Uni-ZAP XR	94	1830	87	1786	128	128	186	1	26	27	45
18	HRACJ35	209878 05/18/98	pCMVSPORT 3.0	28	1863	8	1863	99	99	120	1	24	25	472
18	HAWAZ34	209141 07/09/97	pBluescript SK-	95	1134	472	1132	687	687	187	1			33
19	HTWDE26	209878 05/18/98	pSport1	29	1626	1	1626	68	68	121	1	30	31	167
19	HMHBN40	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	96	1772	69	1772	129	129	188	1	30	31	231
20	HUSIB13	209878 05/18/98	pSport1	30	605	1	605	172	172	122	1	32	33	46
21	HBFAFA02	209877 05/18/98	pSport1	31	931	359	931	46	46	123	1	21	22	108

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
22	H2CBT75	209877 05/18/98	pBluescript SK-	32	1407	1	1407	32	124	1	23	24	60
23	HAGDQ42	209877 05/18/98	Uni-ZAP XR	33	1526	1	1526	126	125	1	18	19	248
24	HBMCI42	209877 05/18/98	pBluescript	34	1737	41	1580	244	126	1	44	45	248
25	HDPBQ71	209877 05/18/98	pCMVSPORT 3.0	35	2242	6	2242	24	127	1	33	34	612
26	HCEJG71	209877 05/18/98	Uni-ZAP XR	36	2235	2	2235	28	128	1	25	26	447
27	HELHL48	209877 05/18/98	Uni-ZAP XR	37	2971	560	2557	629	129	1	16	17	291
27	HSKCT36	209580 01/14/98	Uni-ZAP XR	98	1955	1	1955	31	190	1	18	19	184
28	HISAQ04	209877 05/18/98	pSport1	38	1163	1	1163	61	130	1	21	22	78
29	HJACB89	209877 05/18/98	pBluescript SK-	39	1932	28	1930	95	131	1	23	24	333
30	HTECC05	209877 05/18/98	Uni-ZAP XR	40	881	1	881	27	132	1	15	16	164

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
31	HB3LF01	209877 05/18/98	Uni-ZAP XR	41	1932	201	1931	217	133	1	46	47	244
32	HBXGP60	209877 05/18/98	ZAP Express	42	1164	1	1164	143	134	1	22	23	55
33	HCE5B20	209877 05/18/98	Uni-ZAP XR	43	1105	1	1105	237	135	1	25	26	54
34	HCMSQ56	209877 05/18/98	Uni-ZAP XR	44	1262	1	1262	148	136	1	19	20	88
35	HCNAH57	209877 05/18/98	Lambda ZAP II	45	517	1	517	35	137	1	33	34	61
36	HCUEP91	209877 05/18/98	ZAP Express	46	858	2	858	266	138	1	20	21	105
37	HDPCJ91	209877 05/18/98	pCMVSPORT 3.0	47	6107	1	6107	131	139	1	28	29	51
38	HDPGK25	209877 05/18/98	pCMVSPORT 3.0	48	703	1	703	345	140	1	33	34	119
39	HE2DY70	209877 05/18/98	Uni-ZAP XR	49	639	1	639	137	141	1	45	46	58
40	HE2NV57	209877 05/18/98	Uni-ZAP XR	50	867	1	867	99	142	1	36	37	99

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
41	HETBR16	209877 05/18/98	Uni-ZAP XR	51	1569	1	1569	161	161	143	1	21	22	64
42	HFXDGI3	209877 05/18/98	Lambda ZAP II	52	1196	1	1196	43	43	144	1	37	38	66
43	HFXKY27	209877 05/18/98	Lambda ZAP II	53	945	1	945	44	44	145	1	19	20	58
44	HHPEC09	209877 05/18/98	Uni-ZAP XR	54	488	1	488	71	71	146	1	19	20	55
45	HISAD54	209877 05/18/98	pSport1	55	2860	1	2860	172	172	147	1	19	20	65
46	HJBCY35	209877 05/18/98	pBluescript SK-	56	1559	93	1272	232	232	148	1	23	24	327
47	HKAEA19	209877 05/18/98	pCMVSPORT 2.0	57	2064	1	1909	83	83	149	1	21	22	89
48	HKGDL36	209877 05/18/98	pSport1	58	1050	1	1050	55	55	150	1	33	34	148
49	HLDBS43	209877 05/18/98	pCMVSPORT 3.0	59	2533	1	2533	73	73	151	1	26	27	390
50	HLWAD92	209877 05/18/98	pCMVSPORT 3.0	60	899	1	899	197	197	152	1	34	35	98

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
51	HL YB115	209877 05/18/98	pSport1	61	1079	1	1079	92	92	153	1	22	23	60
52	HMEJE05	209889 05/22/98	Lambda ZAP II	62	1928	1	1928	25	25	154	1	30	31	392
53	HNGIX55	209889 05/22/98	Uni-ZAP XR	63	781	1	781	121	121	155	1	19	20	74
54	HNHEX30	209889 05/22/98	Uni-ZAP XR	64	1194	1	1194	138	138	156	1	15	16	81
55	HPJB133	209889 05/22/98	Uni-ZAP XR	65	1677	1	1677	236	236	157	1	31	32	53
56	HRABA80	209889 05/22/98	pCMVSPORT 3.0	66	1237	1	1237	130	130	158	1	28	29	102
57	HRACD80	209889 05/22/98	pCMVSPORT 3.0	67	1934	1	1934	191	191	159	1	16	17	575
57	HRACD80	209889 05/22/98	pCMVSPORT 3.0	99	1958	1	1958	191	191	191	1	16	17	146
58	HSLCX03	209889 05/22/98	Uni-ZAP XR	68	3300	984	2729	677	677	160	1	22	23	643
58	HSLCX03	209889 05/22/98	Uni-ZAP XR	100	2444	1	2444	392	392	192	1	22	23	124

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
59	HT5GJ57	209889 05/22/98	Uni-ZAP XR	69	1797	92	1797	122	122	161	1	25	26	190
60	HTACS42	209889 05/22/98	Uni-ZAP XR	70	1373	1	1373	213	213	162	1	29	30	63
61	HTEKE40	209889 05/22/98	Uni-ZAP XR	71	1579	1	1579	173	173	163	1	47	48	117
62	HTOBX69	209889 05/22/98	Uni-ZAP XR	72	1028	1	1028	28	28	164	1	20	21	42
63	HUVEO77	209889 05/22/98	Uni-ZAP XR	73	3674	1	3674	55	55	165	1	27	28	47
64	H2CBG48	209889 05/22/98	pBluescript SK-	74	2797	1	2797	125	125	166	1	25	26	45
65	H2CBU83	209889 05/22/98	pBluescript SK-	75	2703	1	2703	157	157	167	1	30	31	207
65	H2CBU83	209889 05/22/98	pBluescript SK-	101	2709	1	2709	157	157	193	1	30	31	51
66	HAPNY94	209889 05/22/98	Uni-ZAP XR	76	742	1	742	94	94	168	1	29	30	50
67	HBJHZ58	209889 05/22/98	Uni-ZAP XR	77	1825	1	1825	102	102	169	1	29	30	42

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
68	HCE2B33	209889 05/22/98	Uni-ZAP XR	78	1674	1	1668	67	67	170	1	18	19	55
69	HDPBQ02	209889 05/22/98	pCMVSPORT 3.0	79	2191	291	2191	460	460	171	1	24	25	108
70	HFIYI70	209889 05/22/98	pSPORT1	80	1335	1	1335	43	43	172	1	15	16	50
71	HDPOZ56	209889 05/22/98	pCMVSPORT 3.0	81	1867	415	1867	103	103	173	1	21	22	566
71	HDPOZ56	209889 05/22/98	pCMVSPORT 3.0	102	1722	1	1722	59	59	194	1	21	22	319

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further

below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed

sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the

information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, supra.) However, the two
5 methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a
10 protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes
15 vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that
20 in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the
25 naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

30 Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

- 5 By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the
- 10 polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire
- 15 sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

- As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known
- 20 computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences
- 25 are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty
- 30 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject
5 sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then
10 subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the
15 purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent
20 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the
25 deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes
30 of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95%

"identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the

query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter

the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced
5 for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an
10 organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA
15 technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., *J. Biol. Chem.* 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after
20 deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., *J. Biotechnology* 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological
25 activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (*J. Biol. Chem.* 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every
30 possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See,

Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or
5 C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular
10 polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions,
15 inversions, repeats, and substitutions selected according to general rules known in the art so as to have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

20 The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these
25 positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function.
30 For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used.

(Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or fragments thereof (e.g., the mature form and/or other fragments described herein), is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

15

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-

30

1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or
5 smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid
10 sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-
15 40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

20 Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the
25 mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

30 Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and

alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

- 5 Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to
10 an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

- 15 In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In
20 contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

- Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985)
25 further described in U.S. Patent No. 4,631,211.)

- In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson
30 et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes
5 the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes
10 in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear
15 more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

20

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the
25 polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present
30 invention include not only heterologous signal sequences, but also other heterologous

functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., *Nature* 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., *J. Biochem.* 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See,

D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide.

- 5 In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein.
- 10 Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

15

Vectors, Host Cells, and Protein Production

- The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral
- 20 vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

- The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate,
- 25 such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

- The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac
- 30 promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The

expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or
5 UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples
10 of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera Sf9* cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-
15 9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1
20 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods
25 are described in many standard laboratory manuals, such as Davis et al., *Basic Methods In Molecular Biology* (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from
30 recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography,

phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

5 Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant,
10 insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine
15 encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

20 In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide
25 sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24,
30 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al.,

Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

5

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

10

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

15

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids

20 containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides

25 can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved

30 using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however,

polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques." Pergamon Press, New York (1988).

5 For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

10 Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins
15 University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined.
20 First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide
25 and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using
30 polynucleotides of the present invention. Any of these alterations (altered expression,

chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods
5 rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J.
10 Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can
15 be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate
20 manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel.
25 In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA
30 markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using
5 this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification
10 techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erich, H., PCR Technology, Freeman and Co.
15 (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

20 There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by
25 organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences
30 in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA

antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

5 Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

 A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M.,
10 et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such
15 as iodine (¹²⁵I, ¹²¹I), carbon (¹⁴C), sulfur (³⁵S), tritium (³H), indium (¹¹²In), and technetium (^{99m}Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

 In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo
20 imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for
25 the relevant hybridoma.

 A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ¹³¹I, ¹¹²In, ^{99m}Tc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or
30 intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety

needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of ^{99m}Tc . The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein.

- 5 In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which
10 involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

- 15 Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to
20 activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can
25 also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as
30 molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also

be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

5

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

10

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

15

20

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency,

25

30

Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

5 Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet
10 disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

15 A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the
20 present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

 Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic
25 anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary
30 Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstrom's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the

present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, Haemophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis,

opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, 5 Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

10 Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and 15 Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these 20 symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the 25 patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

30 A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See,

Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

30

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable

of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell
5 membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

10 The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations,
15 polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

20 Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers.
25 The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

30 Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a

candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) 5 determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as 10 discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be 15 used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive disorders), tendency for violence, 20 tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat 25 content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated 30 nucleic acid molecule comprising a nucleotide sequence which is at least 95%

identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of
5 positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of
10 positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the
15 range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide
20 sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

25 A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X
30 in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under
5 stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which
10 comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide
15 sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of
20 at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

25 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule
30 comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least

one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the
5 Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid
10 sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

15 Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

20 Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

25 Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

30 Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in

the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence
5 at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a
10 polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained
15 in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group
20 consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least
25 one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence
30 selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide

comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino

acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- 5 In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at
10 least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number
15 shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide
20 comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

25 Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

30 Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is

expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of
5 SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1
10 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an
15 isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

20

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector.
25 Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being
30 isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited</u>
	<u>Plasmid</u>	
	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
5	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSPORT 2.0	pCMVSPORT 2.0
	pCMVSPORT 3.0	pCMVSPORT 3.0
10	pCR [®] 2.1	pCR [®] 2.1
	<p>Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Altling-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Altling-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.</p>	
25	<p>Vectors pSport1, pCMVSPORT 2.0 and pCMVSPORT 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain</p>	
30		

XL-1 Blue. Vector pCR[®]2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al.,
5 Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional
10 plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample
15 may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ
20 ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ³²P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular
25 Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection
30 agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for

bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the
5 SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the
3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired
cDNA using the deposited cDNA plasmid as a template. The polymerase chain
reaction is carried out under routine conditions, for instance, in 25 µl of reaction
mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is
10 1.5-5 mM MgCl₂, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25
pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR
(denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1
min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The
amplified product is analyzed by agarose gel electrophoresis and the DNA band with
15 expected molecular weight is excised and purified. The PCR product is verified to be
the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding
portions of a gene which may not be present in the deposited clone. These methods
include but are not limited to, filter probing, clone enrichment using specific probes,
20 and protocols similar or identical to 5' and 3' "RACE" protocols which are well
known in the art. For instance, a method similar to 5' RACE is available for
generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et
al., *Nucleic Acids Res.* 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a
25 population of RNA presumably containing full-length gene RNA transcripts. A
primer set containing a primer specific to the ligated RNA oligonucleotide and a
primer specific to a known sequence of the gene of interest is used to PCR amplify
the 5' portion of the desired full-length gene. This amplified product may then be
sequenced and used to generate the full length gene.

30 This above method starts with total RNA isolated from the desired source,
although poly-A+ RNA can be used. The RNA preparation can then be treated with

phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction
5 leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific
10 to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

15 A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

20 Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After
25 labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H)
30 or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to

manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

5 **Example 4: Chromosomal Mapping of the Polynucleotides**

 An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is
10 repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in
15 the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

 A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA
20 sequence; as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc.,
25 Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

 The pQE-9 vector is digested with BamHI and XbaI and the amplified
30 fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain

M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r).

Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and
5 confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG
10 (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic
15 agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The
20 QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

25 The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4,
30 containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250

mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further
5 includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains:
1) a neomycinphosphotransferase gene as a selection marker, 2) an *E. coli* origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a
10 Shine-Delgarno sequence, and 6) the lactose operon repressor gene (*lacIq*). The origin of replication (*oriC*) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with *NdeI* and
XbaI, *BamHI*, *XhoI*, or *Asp718*, running the restricted product on a gel, and isolating
15 the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for *NdeI* (5' primer) and *XbaI*, *BamHI*, *XhoI*, or *Asp718* (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

20 The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide
25 expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at
15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit
30 weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM

Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate
5 is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the
10 pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA
15 by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The
20 filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4
25 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40
30 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium

acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

5 The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 μ g of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

10

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector
15 contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak
20 *Drosophila* promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such
25 as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

30 Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in

Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in
5 Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment
10 then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a
15 commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified
20 by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™
25 baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are
30 added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded

in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

- 5 After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell
- 10 culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded
- 15 in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

- To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of
- 20 infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of ³⁵S-methionine and 5 µCi ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation.
- 25 The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

- 30 **Example 8: Expression of a Polypeptide in Mammalian Cells**

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIV and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human HeLa, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QCI-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a

chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession
5 No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the
10 polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

15 A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

20 The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and
25 purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for
30 transfection. Five μ g of the expression plasmid pC6 is cotransfected with 0.5 μ g of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo

contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

15 **Example 9: Protein Fusions**

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

30 Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These

primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

```

GGGATCCGGAGCCCAAATCTTCTGACAAAACCTCACACATGCCCCACCGTGC
CCAGCACCTGAATTTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAA
CCCAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGT
GGTGGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGG
ACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTA
CAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACT
GGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA
ACCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
CACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAG
GTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGT
GGAGTGGGAGAGCAATGGGCAGCCGAGAACAACTACAAGACCACGCCT
CCCGTGCTGGACTCCGACGGCTCCTTCTCCTCTACAGCAAGCTCACCGTG
GACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCA
TGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG
GTAAATGAGTGCGACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

```

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the
5 production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are
10 monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures
15 involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about
20 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are
25 selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can
30 be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is

possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody
5 whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')₂ and other fragments of the
10 antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic
15 chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in
20 the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

25

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in
30 Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2×10^5 cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media. either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl₂ (anhyd); 0.00130 mg/L CuSO₄·5H₂O; 0.050 mg/L of Fe(NO₃)₃·9H₂O; 0.417 mg/L of FeSO₄·7H₂O; 311.80 mg/L of KCl; 28.64 mg/L of MgCl₂; 48.84 mg/L of MgSO₄; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO₃; 62.50 mg/L of NaH₂PO₄·H₂O; 71.02 mg/L of Na₂HPO₄; 4320 mg/L of ZnSO₄·7H₂O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H₂O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H₂O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H₂O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H₂O; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B₁₂; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

- 5 On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the
10 polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

15 **Example 12: Construction of GAS Reporter Construct**

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The
20 binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in
25 many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus
30 upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2,

Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

	<u>Ligand</u>	<u>tyk2</u>	<u>JAKs</u>		<u>STATs</u>	<u>GAS(elements) or ISRE</u>
			<u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>	
	<u>IFN family</u>					
5	IFN-a/B	+	+	-	-	1,2,3 ISRE
	IFN-g		+	+	-	1 GAS (IRF1>Lys6>IFP)
	IL-10	+	?	?	-	1,3
	<u>gp130 family</u>					
10	IL-6 (Pleiotrophic)	+	+	+	?	1,3 GAS (IRF1>Lys6>IFP)
	IL-11(Pleiotrophic)	?	+	?	?	1,3
	OnM(Pleiotrophic)	?	+	+	?	1,3
	LIF(Pleiotrophic)	?	+	+	?	1,3
	CNTF(Pleiotrophic)	-/+	+	+	?	1,3
15	G-CSF(Pleiotrophic)	?	+	?	?	1,3
	IL-12(Pleiotrophic)	+	-	+	+	1,3
	<u>g-C family</u>					
	IL-2 (lymphocytes)	-	+	-	+	1,3,5 GAS
20	IL-4 (lymph/myeloid)	-	+	-	+	6 GAS (IRF1 = IFP >>Ly6)(IgH)
	IL-7 (lymphocytes)	-	+	-	+	5 GAS
	IL-9 (lymphocytes)	-	+	-	+	5 GAS
	IL-13 (lymphocyte)	-	+	?	?	6 GAS
	IL-15	?	+	?	+	5 GAS
25	<u>gp140 family</u>					
	IL-3 (myeloid)	-	-	+	-	5 GAS (IRF1>IFP>>Ly6)
	IL-5 (myeloid)	-	-	+	-	5 GAS
	GM-CSF (myeloid)	-	-	+	-	5 GAS
30	<u>Growth hormone family</u>					
	GH	?	-	+	-	5
	PRL	?	+/-	+	-	1,3,5
	EPO	?	-	+	-	5 GAS(B-CAS>IRF1=IFP>>Ly6)
35	<u>Receptor Tyrosine Kinases</u>					
	EGF	?	+	+	-	1,3 GAS (IRF1)
	PDGF	?	+	+	-	1,3
	CSF-1	?	+	+	-	1,3 GAS (not IRF1)
40						

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

10 5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCC
GAAATGATTTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

15 PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

20 5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAA
TGATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCG
CCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCT
CCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCC
TCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCT
25 AGGCTTTTGCAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter
30 molecules that can be used instead of SEAP include chloramphenicol

acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and
5 XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-
10 SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into
15 mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and
20 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

25

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12.
30 Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells

(ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-
5 SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is
10 demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1% Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life
15 Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final concentration of 10^7 cells/ml. Then add 1ml of 1×10^7 cells in OPTI-MEM to T25
20 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Gentamicin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in
25 Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100
30 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

The above protocol may be used in the generation of both transient, as well as, stable transfected cells, which would be apparent to those of skill in the art.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell

Growth & Differentiation, 5:259-265) is used. First, harvest 2×10^7 U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

- 5 Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 1 mM MgCl_2 , and 675 uM CaCl_2 . Incubate at 37°C for 45 min.

- 10 Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

- 15 These cells are tested by harvesting 1×10^8 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^5 cells/ml. Plate 200 ul cells per well in the 96-well plate (or 1×10^5 cells/well).

- 20 Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

- 25 When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat pheochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor).

- 5 The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene
10 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCCAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

- Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified
15 product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

- To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30
20 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

- PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-
25 inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

- Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine
30 protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine

growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS
5 (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as 5×10^5
10 cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold
15 induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

NF- κ B (Nuclear Factor κ B) is a transcription factor activated by a wide
20 variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- κ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- κ B appears to shield cells from apoptosis), B and T-cell development,
25 anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κ B is retained in the cytoplasm with I- κ B (Inhibitor κ B). However, upon stimulation, I- κ B is phosphorylated and degraded, causing NF- κ B to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κ B include IL-2, IL-6, GM-CSF, ICAM-1 and
30 class I MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- κ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- κ B would be useful in treating diseases. For example, inhibitors of NF- κ B could be used to treat those
 5 diseases related to the acute or chronic activation of NF- κ B, such as rheumatoid arthritis.

To construct a vector containing the NF- κ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- κ B binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary
 10 to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:
 5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC
 TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:
 15 5':GCGGCAAGCTTTTGGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)
 Sequencing with the T7 and T3 primers confirms the insert contains the following
 20 sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGACTTTCC
 ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCC
 ATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGA
 25 CTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTA
 TTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAA
 GCTT:3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-
 30 promoter plasmid (Clontech) with this NF- κ B/SV40 fragment using XhoI and

HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- κ B/SV40/SEAP cassette is removed from the above NF- κ B/SEAP vector using restriction enzymes
5 SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF- κ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF- κ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly,
10 the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

15 As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 μ l of 2.5x
20 dilution buffer into Optiplates containing 35 μ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 μ l Assay Buffer and incubate at room
25 temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 μ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

30 Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small**5 Molecule Concentration and Membrane Permeability**

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants

which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

- 5 The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-4 (Molecular Probes, Inc.; catalog no. F-14202), used here.

- 10 For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

- 15 A stock solution of 1 mg/ml fluo-4 is made in 10% pluronic acid DMSO. To load the cells with fluo-4, 50 ul of 12 ug/ml fluo-4 is added to each well. The plate is incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

- 20 For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-4 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

25 For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-4. The supernatant is added to the well, and a change in fluorescence is detected.

- 30 To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm;

and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca^{++} concentration.

5 **Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity**

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth
10 factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor
15 dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily
20 of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins
25 capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr
30 with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St.

Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na₃VO₄, 2 mM Na₄P₂O₇ and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂⁺ (5mM ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

10 The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mM EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide.

15 Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and

20 incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

25 **Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity**

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be

30 used. For example, as described below one particular assay can detect tyrosine

phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv.

et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

10 A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

15 For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

20 The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbound polypeptide.

25 Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbound conjugate.

30 Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard

curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

5

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 $\mu\text{g/kg/day}$ to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day , and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 $\mu\text{g/kg/hour}$ to about 50 $\mu\text{g/kg/hour}$, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's

solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical

compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and
5 separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS,
10 penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

15 pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified
20 using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions
25 appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with
30 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector.

The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, 5 containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. 10 If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or 15 after having been grown to confluence on cytodex 3 microcarrier beads.

Example 27: Method of Treatment Using Gene Therapy - In Vivo

20 Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter 25 or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) Cardiovasc. Res. 35(3):470-479, Chao J et al. (1997) Pharmacol. Res. 35(6):517-522, Wolff J.A. (1997) Neuromuscul. Disord. 30 7(5):314-318, Schwartz B. et al. (1996) Gene Ther. 3(5):405-411, Tsurumi Y. et al. (1996) Circulation 94(12):3281-3290 (incorporated herein by reference).

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection

into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 μ m cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that 5 quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be used to extrapolate proper dosages and other treatment parameters in humans and other animals using naked 10 DNA.

Example 28: Transgenic Animals.

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, 15 guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

20 Any technique known in the art may be used to introduce the transgene (i.e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994); Carver et al., Biotechnology (NY) 11:1263-1270 (1993); Wright et al., Biotechnology (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus 25 mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the 30 polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., Science 259:1745 (1993)); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-

mediated gene transfer (Lavitano et al., Cell 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones
5 containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

The present invention provides for transgenic animals that carry the transgene
10 in all their cells, as well as animals which carry the transgene in some, but not all their cells, *i.e.*, mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et
15 al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is
20 to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only
25 that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the
30 recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA

expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also
5 be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more
10 than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for
15 screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not
20 limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

25 **Example 29: Knock-Out Animals.**

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (*E.g.*, see Smithies et al., *Nature* 317:230-234 (1985); Thomas & Capecchi, *Cell* 51:503-512 (1987); Thompson et al., *Cell* 5:313-321 (1989); each of which is incorporated by
30 reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding

regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e.g., see Thomas & Capecchi 1987 and Thompson 1989, *supra*). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e.g., knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and
5 Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For
10 example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological
15 function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

It will be clear that the invention may be practiced otherwise than as
20 particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other
25 disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>178</u> . line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 18, 1998</u>	Accession Number <u>209878</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-top: 1px solid black; border-bottom: 1px solid black;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="" type="checkbox"/> This sheet was received with the international application</div> <div style="padding: 5px;">Authorized officer <u>Missy Walker</u> <u>International Division</u> <u>703-205-6032</u> <u>missy.walker@usdoj.gov</u></div>	<div style="text-align: center; border-top: 1px solid black; border-bottom: 1px solid black;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="padding: 5px;">Authorized officer</div>
---	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>179</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>February 26, 1997</u>	Accession Number <u>97898</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit") 	
For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer <u>Misty Walker</u> <u>International Division</u> <u>705-615-3632</u> <u>Washington, D.C.</u>	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>179</u> . line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209044</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit") 	
For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer <u>[Signature]</u> <u>709-000000</u>	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>179</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit August 28, 1997	Accession Number 209225
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable) The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<input checked="" type="checkbox"/> For receiving Office use only This sheet was received with the international application Authorized officer 703-615-2892	<input type="checkbox"/> For International Bureau use only This sheet was received by the International Bureau on: Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>180</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit March 7, 1997	Accession Number 97923
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer	<i>Misty Walker</i> International Division 708-595-2000

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>180</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 22, 1997</u>	Accession Number <u>209071</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit") 	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <u>Patricia A. Schaefer</u> <u>Examination Division</u> <u>705-405-0052</u> <u>microorganisms@uspto.gov</u>	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>180</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit February 12, 1998	Accession Number 209626
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable) The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer:	<i>[Signature]</i> International Division 705-400-6660

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer:	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit March 20, 1998	Accession Number 209683
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<input checked="" type="checkbox"/> For receiving Office use only This sheet was received with the international application	<input type="checkbox"/> For International Bureau use only This sheet was received by the International Bureau on:
Authorized officer [Signature] [Stamp: Division]	Authorized officer

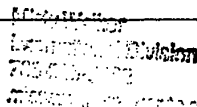
INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit February 25, 1998	Accession Number 209641
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable) The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<input checked="" type="checkbox"/> For receiving Office use only This sheet was received with the international application	<input type="checkbox"/> For International Bureau use only This sheet was received by the International Bureau on:
Authorized officer <i>[Signature]</i> 7/20/99 [Stamp]	Authorized officer

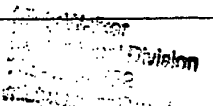
INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit July 9, 1997	Accession Number 209141
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable) The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

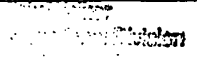
INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit February 26, 1997	Accession Number 97901
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209047</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit") 	
For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer 	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 18, 1998</u>	Accession Number <u>209877</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit") 	

For receiving Office use only	
<input checked="" type="checkbox"/>	This sheet was received with the international application
Authorized officer	<u>Sally Walker</u> <u>International Division</u> <u>703-605-6832</u>

For International Bureau use only	
<input type="checkbox"/>	This sheet was received by the International Bureau on:
Authorized officer	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>182</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>January 14, 1998</u>	Accession Number <u>209580</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit") 	

For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer 	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer
--	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>185</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 22, 1998</u>	Accession Number <u>209889</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable) The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<div>For receiving Office use only</div> <div><input checked="" type="checkbox"/> This sheet was received with the international application</div> <div>Authorized officer <u>[Signature]</u></div>	<div>For International Bureau use only</div> <div><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div>Authorized officer</div>

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
- 5 (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a
10 polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - 15 (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X,
20 having biological activity;
 - (f) a polynucleotide which is a variant of SEQ ID NO:X;
 - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
 - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
 - 25 (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
- 5 3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- 10 4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- 15 5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 20 6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.
- 25 8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.
9. A recombinant host cell produced by the method of claim 8.
- 30 10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

5 (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

(c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

10 (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

15 (g) a variant of SEQ ID NO:Y;

(h) an allelic variant of SEQ ID NO:Y; or

(i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

20

13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

25 14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15. A method of making an isolated polypeptide comprising:

30 (a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and

(b) recovering said polypeptide.

16. The polypeptide produced by claim 15.
17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount
5 of the polypeptide of claim 11 or the polynucleotide of claim 1.
18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
- (a) determining the presence or absence of a mutation in the polynucleotide of
10 claim 1; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.
19. A method of diagnosing a pathological condition or a susceptibility to
15 a pathological condition in a subject comprising:
- (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological
20 condition based on the presence or amount of expression of the polypeptide.
20. A method for identifying a binding partner to the polypeptide of claim
11 comprising:
- (a) contacting the polypeptide of claim 11 with a binding partner; and
- (b) determining whether the binding partner effects an activity of the
25 polypeptide.
21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.
22. A method of identifying an activity in a biological assay, wherein the
30 method comprises:
- (a) expressing SEQ ID NO:X in a cell;
- (b) isolating the supernatant:

- (c) detecting an activity in a biological assay; and
- (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 20.

1

Sequence Listing

<110> Human Genome Sciences, Inc., et al.

<120> 71 Human Secreted Proteins

<130> PZ030PCT

<140> Unassigned

<141> 1999-07-14

<150> 60/092,956

<151> 1998-07-15

<150> 60/092,921

<151> 1998-07-15

<150> 60/092,922

<151> 1998-07-15

<160> 262

<170> PatentIn Ver. 2.0

<210> 1

<211> 733

<212> DNA

<213> Homo sapiens

<400> 1

gggatccgga gcccaaatct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg	60
aattcgaggg tgcaccgtca gtcttctctt tccccccaaa acccaaggac accctcatga	120
tctcccgac tcctgaggtc acatgcgtgg tggtaggacgt aagccacgaa gacctgagg	180
tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg	240
aggagcagta caacagcacg tacctgtggt tcagcgtcct caccgtcctg caccaggact	300
ggctgaatgg caaggagtac aagtgcagg tctccaaaca agccctccca acccccatcg	360
agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc	420
catcccgga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct	480
atccaagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga	540
ccacgcctcc cgtgctggac tccgacggct ccttcttct ctacagcaag ctcaccgtgg	600
acaagagcag gtggcagcag gggaacgtct tctcatgtc cgtgatgcat gaggtcttgc	660
acaaccacta cacgcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc	720
gactctagag gat	733

<210> 2

<211> 5

<212> PRT

<213> Homo sapiens

<220>

<221> Site

<222> (3)

<223> Xaa equals any of the twenty naturally occurring L-amino acids

<400> 2

Trp Ser Xaa Trp Ser

1

5

<210> 3
<211> 86
<212> DNA
<213> Homo sapiens

<400> 3
gcgcctcgag atttccccga aatctagatt tccccgaaat gatttccccg aaatgatttc 60
ccccgaaatat ctgccatctc aattag 86

<210> 4
<211> 27
<212> DNA
<213> Homo sapiens

<400> 4
gcggcaagct ttttgcaaag cctaggc 27

<210> 5
<211> 271
<212> DNA
<213> Homo sapiens

<400> 5
ctcgagattt cccccgaaatc tagatttccc cgaaatgatt tccccgaaat gatttccccg 60
aaatatctgc catctcaatt agtcagcaac catagtcccc ccctaactc cgcccatccc 120
gcccctaact ccgcccagtt ccgcccattc tccgcccatt ggctgactaa ttttttttat 180
ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 240
ttttggaggc ctaggctttt gcaaaaagct t 271

<210> 6
<211> 32
<212> DNA
<213> Homo sapiens

<400> 6
gcgctcgagg gatgacagcg atagaacccc gg 32

<210> 7
<211> 31
<212> DNA
<213> Homo sapiens

<400> 7
gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8
<211> 12
<212> DNA
<213> Homo sapiens

<400> 8
ggggactttc cc 12

<210> 9
<211> 73
<212> DNA
<213> Homo sapiens

<400> 9
gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg 60
ccatctcaat tag 73

<210> 10
<211> 256
<212> DNA
<213> Homo sapiens

<400> 10
ctcgagggga ctttcccggg gactttccgg ggactttccg ggactttcca tctgccatct 60
caattagtca gcaaccatag tcccggccct aactccgccc atcccgcccc taactccgcc 120
cagttccgcc cattctccgc cccatggctg actaattttt ttattttatg cagaggccga 180
ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg 240
cttttgcaaa aagctt 256

<210> 11
<211> 1113
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (393)
<223> n equals a,t,g, or c

<400> 11
gatgctcctt tagcttggag gagtttgta ttaccacact tctgaagcct acttctgtca 60
attcatccaa ctcatctca gtccagtttt gtttccttgc tggtagaggag ttgtgatcct 120
ttggaggaga agaggcattc tggtttttgg aatttttagc cattttgctc tggtttcttc 180
ccatctttgt ggatttatct acctttcacc ttcaatgta gtgacctatg gatggggctct 240
ctgagtggat gtgctcttcc tttctgtttg twagttttct ttctaacagt tagccctct 300
gctgtaggtc tgctggaktt tgctggaggt ccactccaga cctgttttgc ctgggtatca 360
ccagtggagg ctgcagaaca gcaaagattg ctncctgttc tttcctctgg aagcttcgtc 420
tcagagggca cctgccagat gccagccaga gctctcctgt atgaggtgtc tggtagccca 480
tactgggaga ttcctcccag tcaggatata aggaggtcag ggacctactt gaggaggcag 540
tctgaccctt agcagaggtt gaacactgtg ctaggaggtc ctctgctctt ttcagagctg 600
tcaggcgggg cgtataagtc tgctgaatct gtgtccgcag ccacccttc cccaggtgc 660
tctgtcccag ggagatgggg gttttatatt taagtccca actggggctg ctgccttttt 720
ttcagagatg ccctgccag agaggagaaa tctggcagtc tggcctcaga ggccttctg 780
agctgccgtt ggctccaccc agttcaaaact tcccaagggg ctttgtttat actgtaaggg 840
gaaaaccgcc tactcgagcc tcatcaatgg cagacacccc tccccgcgcc aagcttaagt 900
gtcccaggtt gatctcagac tgctgctgtg ctggcagtgga gaatttcaag ccagtggatc 960
ttagtttctg gtgctctgtg ggggtgggac ccattgaacc agactactcg gctccctggc 1020
ttcagccccc tttccaggag agtgaagggt tctgtctcat tggcattcca ggagtcactg 1080
gtgtatggaa aaaaaaaaaa aaagggcggc cgc 1113

<210> 12

<211> 983
<212> DNA
<213> Homo sapiens

<400> 12
ggcacgaggg cagctgcaga gctccagggt tctctgcca caagggcagg ggctgccctt 60
cgcccaggat gactctgctt tccagagcct tggcctccct gggggtggga gtgtggggga 120
tgctaagggt aaatcagggt acagtaagtt gtgggggcag cagggtggagc agcagagtgg 180
cactgggagc tttctcttgg gtgtgcggtg tggccttggt tctgcagcca tcaggtgggg 240
gcttgggact gacttctcct tctgaaggat gctgggaagg tgagctggct ttggcagtgc 300
ttagagctcc ggggggttcc cctcctaga acatgcaagc tctcacaccg gtgcgtcatc 360
atcacaccca tcatcaagcc cacagtggta tactgaacac ctgcccaca aagacggtgg 420
actgctctca gaggagcccc atgaaccacc gatggttaca actatccaat gcctgatggc 480
agacagccag gccaacctcg gcttccactc tctcttctc accctacaat cagccaaagt 540
gacctgagtc atgtagtgtg aagttgcttt ctgctttctc ttgtttgtgc ttttgcgtgt 600
tcttctgccc catactttgt taactccatg agttaaatgc taccattttt cccagacaag 660
tgctgcttct gcaaggaaac ccttctgat ccccccacta tctgaaaagt acctctccag 720
cttgccttct cagggtgctg agcgttctct cccagcctgt catcaccttc ctccatacgc 780
tatggtgtgt tctgtcttc tctagtcttg tcttctttt tctgttagat tgtagctcct 840
tgctgacagg aaccacgcct gctccagctt catacctccc actgctacag cacagaacct 900
gcttctcaga cttacagcaa atgtttgttt gctgaatgaa ttaattaaag ataaagcaaa 960
aaaaaaaaa aaaaaaactc gag 983

<210> 13
<211> 973
<212> DNA
<213> Homo sapiens

<400> 13
ggcacgagcc cagcggaagc caagccacca ggccccccag cgtccacgcg gagcatgaac 60
attgaggatg gcgcgtgccc gcggctcccc gtgcccccg ctgccgccg gtaggatgtc 120
ctggccccac ggggcattgc tcttctctg gctcttctc ccaccctgg gggccggtgg 180
aggtggagtg gccgtgacgt ctgccgccg agggggctcc ccgccggcca cctctgccc 240
cgtggcctgc tctgcagca accaggccag ccgggtgatc tgcacacgga gagacctggc 300
cgaggtccca gccagcatcc cggtaaacac gcggtacctg aacctgcaag agaacggcat 360
ccaggtgatc cggacggaca cgttaagca cctgcccac ctggagattc tgcagctgag 420
caagaacctg gtgcgcaaga tcgaggtggg cgccttcaac gggctgccc gcctcaaac 480
gctggagctt tttgacaacc ggctgaccac ggtgcccac caggccttc agtacctgtc 540
caagctgcgg gagctctggc tgcggaacaa cccatcgag agcatccct cctacgcctt 600
caaccgcgtg cctcgtctgc ggcgcctgga cctgggcgag ctcaagcggc tggatacat 660
ctcggaggcg gccttcgagg ggctggtcaa cctgcgctac ctcaacctgg gcatgtgcaa 720
cctcaaggac atcccccaacc tgacggccct ggtgcgcctg gaggagctgg agctgtcggg 780
caaccggctg gacctgatcc gcccgggctc ctccagggt ctcaccagcc tgcgcaagct 840
gtggctcatg cagccccagg tagccaccat cgagcgcaac gccttcgacg acctcaagtc 900
gctggaggag ctcaacctgt cccacaacaa cctgatgtcg ctgccccacg acctcttcac 960
gcccctgcac cgc 973

<210> 14
<211> 1458
<212> DNA
<213> Homo sapiens

<400> 14
ccacgcgtcc gggaattttc aaaagatcca aacagagact tctgcatct tctgcctttc 60
caacagaagc ggtgatctgc taagtatgag cctgtggctt cctttgtgca tttgagcatg 120
ctgtaattaa gatgagatca gtttcttaga aaaagctttc ctgaatccct ctgacgttgc 180

ctgggatctt	tctgttgatt	cgtcttttct	ggagattggg	acagagcatc	tgtgggtccag	240
ggaagttagt	cctctggcct	caattctgtt	gtggatgtgc	agtgataaagc	gggcattgcg	300
tgccctcggg	gatgcctagt	tcgtggcttc	ctggctgttt	tgctccttctg	tgtctttag	360
ctgtagggtg	ccagctcagg	gagtgggggt	ttggcgcggt	ttccgcgggt	ggcctccttg	420
ctttgccgca	cctccagggt	ctgggcatga	gaggccgtgg	cctcatttct	ggtggataac	480
cttttttagtt	taatagcatc	tttaattaga	tcacagcatt	gaattcaaaa	tttcttctgc	540
aaagaaagt	gtggggcata	agacaccggg	aatgaggag	gaggaagaca	gttgtgtttt	600
ctctttaaac	cttgagctct	agccgatgca	tttgtcagga	aatacagcac	tttgtcttaa	660
gaaaaacaag	aaggaggcgg	ggcgagtggt	ctcacgcctg	taatcccagc	actttgggag	720
gccgaggcgg	gaggatcacc	tgaggtgggg	agtatgagac	caccctgact	aacatggaga	780
gacctgtct	ctactaaaag	tacagaatta	gccgggctgt	gttgcgcatg	cccataatcc	840
cagctactga	ggagacttga	ggtaggagaa	tcacttgaac	ctcagcgcg	gaggttgag	900
tgagtcgaga	tcgcgccagt	gcactccagc	ctgggcaaga	agagcgaaac	tgggtctcaa	960
gttaaaaaaa	gaaagcaagg	aaagagtaat	ttacaacgaa	ggaaaaaac	ccacagcaca	1020
cccttcgagg	ctgtcagcgc	tctcctgatg	tcacagtggc	tgctgtctct	tgggggtggg	1080
gaggtgtggg	gagccagacc	cctggccctg	cctcccgcgc	cccgcctccc	ttctctctct	1140
tactcgggta	agccatagcg	aggcctccgc	tcgtttcaga	tatgaatttg	ttttatagat	1200
tataaatatg	catatacagt	gtatgtataa	agcagaatgc	ctgcctttcc	tggttatttt	1260
ttgtaccata	ttgtaaatga	tattatttat	tctttaccaa	ttttgggaat	aaaaggtgtt	1320
ttggttattt	aatataataa	gagctgttaa	acttctgttt	aaatttccag	ttcaacttgt	1380
aaatgttttt	attgtgcata	aatacatact	aatgttgatc	taaaaaaaa	aaaaaaaaaa	1440
aaaaaaaaaa	aaaaaaaaaa					1458

<210> 15

<211> 2005

<212> DNA

<213> Homo sapiens

<400> 15

ggttgctggc	ccaggtgagc	gggcgcgctg	gtccaggtga	gcggggcgct	ccccgcgacg	60
gcgctgcctg	cccaggcggg	ttcacgtaaa	gacagcgaga	tcctgagggc	cagcccgga	120
ggaggcgtgg	atatggagct	ggctgctgcc	aagtcggggg	cccgcgcgcg	tgccatagcgc	180
gtcctgggga	ctctgtgggg	acgcgcggcg	cgccgcggct	cggggacccg	tagagcccg	240
cgctgcgcgc	atggccctgc	tctcgcgcgc	cgcgctcacc	ctcctgctcc	tcctcatggc	300
cgctgttgtc	aggtgccagg	agcaggccca	gaccaccgac	tggagagcca	ccctgaagac	360
catccggaac	ggcgttcata	agatagacac	gtacctgaac	gccgccttgg	acctcctggg	420
aggcgaggac	ggtctctgcc	agtataaatg	cagtgaagga	tctaagcctt	tcccaggtta	480
tggttataaa	ccctccccac	cgaatggatg	tggctctcca	ctgtttgggt	ktcatcttaa	540
cattggtatc	ccttccttga	caaagtgttg	caaccaacac	gacaggtgct	atgaracctg	600
tggcaaaagc	aagaatgact	gtgatgaaga	attccagtat	tgctcttcca	agatctgccg	660
agatgtacag	aaaacactag	gactaactca	gcatgttcag	gcatgtgaaa	caacagtgga	720
gctcttgttt	gacagtggtta	tacatttagg	ttgtaaacca	tatctggaca	gccaacgagc	780
cgcctgcagg	tgctcattatg	aagaaaaaac	tgatctttta	aggagatgcc	gacagctagt	840
gacagatgaa	gatggaagaa	cataaccttt	gacaaataac	taatgttttt	acaacataaa	900
actgtcttat	ttttgtgaaa	ggattatttt	gagaccttaa	aataattttat	atcttgatgt	960
taaaacctca	aagcaaaaaa	agtgggggag	atagtggagg	gagggcacgc	ttgtcttctc	1020
aggatcttct	cccagcattg	ctcccttact	tagtatgcca	aatgtcttga	ccaatatcaa	1080
aaacaagtgc	ttgttttagcg	gagaattttg	aaaagaggaa	tatataactc	aattttcaca	1140
accacattta	ccaaaaaag	agatcaaata	taaaattcat	cataatgtct	gttcaacatt	1200
atcttatttg	gaaaatgggg	aaattatcac	ttacaagtat	ttgtttacta	tgaattttta	1260
aatacacatt	tatgcctaga	aggaacggac	tttttttttc	tattttaatt	acacataata	1320
tgtaattaaa	gtmcaacata	atatgttgtt	tctctgtagc	ccgttgagca	tatgagtaag	1380
tcacatttct	attaggacta	cttmcaagga	caaggtttcc	atttttccag	ttgtaaaatt	1440
ggaaccatca	gctgataacc	tcgtaggggg	caaccccgag	atagctaagt	gttatgtaat	1500
atgcctagaa	ggtgatgtga	atgcgattca	gaagcatagc	cactcccat	ttatgagcta	1560
ctcacatgac	aaatgtcatc	ttttgtcata	accttttgca	agttagagaa	aagatggatt	1620
taatgagata	aatgaaaaga	tatttamcct	aatatatcaa	ggcactattt	gctgttatgc	1680

6

tttgttattt	atttcccagc	acttgttcct	tattgtatag	tttttaaaga	ctgtaacctt	1740
ttactaactg	tgggtcttact	aaaattttgt	cttgatactg	cttttcaaaa	agcctttaat	1800
tagagccaaa	aggatggaaa	aggcaagata	taaatgcctt	ttatagatct	cttatattaca	1860
ttgaaaatta	ttaccatatt	tttagagcaa	atccaagaaa	acttcaacag	cttctgaaga	1920
tgtctatgaa	tgttgaaaac	ttttcaatst	cttggratgc	tcakttaatt	cgcagaccgg	1980
cttaacggat	taaacgcccc	cccc				2005

<210> 16

<211> 943

<212> DNA

<213> Homo sapiens

<400> 16

ggcacgagct	cgcgccggcc	cgcaggggct	ctccccggag	gctcagcccc	ctctgctccc	60
catgggcaac	tgccaggcag	ggcacaacct	gcacctgtgt	ctggccacc	acccacctct	120
ggtctgtgcc	actttgatcc	tgctgtcct	tgccctctct	ggcctgggccc	ttggcagctt	180
cctcctcacc	cacaggactg	gcctgcgcac	cctgacatcc	cccaggactg	ggtctctttt	240
ttgagatctt	ttggccagct	gaccctgtgt	cccaggaatg	ggacagtcac	agggaagtgg	300
cgagggtctc	acgtcgtggg	cttgcctgacc	accttgaact	tcggagacgg	tccagacagg	360
aacaagaccc	ggacattcca	ggccacagtc	ctgggaagtc	agatgggatt	gaaaggatct	420
tctgcaggac	aactggctct	tatcacagcc	agggtgacca	cagaaaggac	tgagggaacc	480
tgcttatatt	ttagtgctgt	tccaggaatc	ctaccctcca	gccagccacc	catatcctgc	540
tcagaggagg	gggctggaaa	tgccacctg	agccctagaa	tggttgaggga	atgtgttagt	600
gtctggagcc	atgaaggcct	tgtgctgacc	aagctgctca	cctcggaggga	gctggctctg	660
tgtggctcca	ggctgctggt	cttgggctcc	ttcctgcttc	tcttctgtgg	ccttctctgc	720
tgtgtcactg	ctatgtgctt	ccaccgcgc	cgggagctcc	actggtctag	aacccgctc	780
tgagggcact	ggcctagtct	ccgacttggt	tctcaggtgt	gaatcaactt	cttgggcctt	840
ggctctgagt	tggaaaaggt	tttagaaaaa	gtgaagagct	ggaatgtggg	ggaaaaataa	900
aagctttttt	gccccaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaa		943

<210> 17

<211> 1503

<212> DNA

<213> Homo sapiens

<400> 17

cagggttcctc	tcagtamarc	ctcarsecga	ggttcccttc	ctcttgcatc	catgtgtgtg	60
tttcaraggc	ggccatcctt	ccctacttcc	agatccttgt	agggcagttg	gtggagggtg	120
ggaggcacc	cgggtgtgccc	tccatgaagc	cctgtgccag	tcactgggct	gcaaggctga	180
ggaaattgtg	tccgtgtcag	aaagctcctc	agctcagagg	tgctgggtacc	tcctgcgtgg	240
taggaaggca	gggggaagag	gccctgcttc	tcctgttctc	tttgccctta	tgagacttga	300
gagtctgtgt	catctgtgcc	ttgcatgtct	ttttttcaga	ctccctgcga	caaggactgt	360
gtactgcatg	aatgaggctg	agatagttag	tggtgctctg	ggaatcctga	ttgagagccg	420
yaaacaggam	aaggcctgcg	agcagccggc	cctggcgggg	gctgataacc	cagagcactc	480
ccctccctgc	tccgtgtcgc	ctcacacaag	ttctgggagc	agcagttagg	aagaggacag	540
tgggaaaacag	gcaactgrctc	caggcctcag	cccttcccag	aggccggggg	gttccagctc	600
tgctgttagc	aggagccctg	aggaggagga	ggaagaggat	gtgctgaaat	acgtccggga	660
gatctttttc	agctagggca	taaaactgtc	actgaactgt	ctgccgagag	cagctggagg	720
acagctgagc	ttccactggt	gctgctgggc	cgcccgctct	tgggaaatggg	gctctctgtg	780
ctcctacctt	tgtgccttct	tgggcctggc	agattcacct	caggccagaa	gcccctggac	840
actccgggcc	ttggggctgc	cgttctgagt	gtgcggaagg	caggactcaa	aatgagatcc	900
catttgactc	cctctgtatg	tactgtgccc	tctcctggct	cttgaggctc	tgaggtccca	960
attgtctgtg	ttagtcagtg	accagggtcc	agggaaaatg	atgtcatgtg	gtggtccaac	1020
ttactggaac	caaagagaca	gtacttttga	aagaaaaggga	tcactgccag	gtgcactgga	1080
attgctacag	tttagtccgc	atgatctctc	ctgaaggagg	aagcctgttt	caaaaatagt	1140
ttccatcatg	agtctatcaa	tgagctccca	cctctccagc	cagcctagaa	agcaaacgag	1200

7

ctgcccacag	ttctctgccc	tgtctgggag	gttgaggcca	cagtgtatag	actggtaagc	1260
cagacaggcc	tcctcccgcg	agctgctacc	ttgctttcac	ctgtaccttg	gtccccgggc	1320
agctagctat	aaagcaagag	ggacaggagc	ccagaagaga	cactgaggac	aagagatcac	1380
accagagtac	atgtctctgc	ctctgttttc	agtgtggctt	tggacaggaa	tatatgaata	1440
aatcactgcc	atacaggttt	tccaatacac	aagtgctaga	aaatacacac	aattccccaa	1500
tga						1503

<210> 18
 <211> 1512
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (207)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (209)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (521)
 <223> n equals a,t,g, or c

<400> 18							
gcagagcccc	tgggtgtgag	aagctcgtct	cccggtgggtt	gcattggctc	tgccctatct		60
ctgcctccag	cacccagggc	ggccgcagat	ggcagtgtct	ctggggacag	cagctgcgaa		120
tgagtccacg	ggccaacgct	gagctgctca	ggctgaggcg	gtgtgctcag	cacagagccc		180
ccggaactgg	catctgcagg	gcgtgancna	aggccgccgc	gatgccgcac	ttcctggact		240
ggttcgtgcc	ggtctacttg	gtcatctcgg	tcctcattct	ggtgggcttc	ggcgccctga		300
tctactactt	cgagccgggc	ctgcaggagg	cgcacaagtg	gcgcatgcag	cgccccctgg		360
tggaccgcsa	cctccgcaag	acgctaattg	tgcgcgacaa	cctggccttc	ggcggccccg		420
aggtctgagc	cgacttgcaa	aggggatagg	csggcggcac	cgggcgccct	ccccagcccc		480
gccccgcccc	cccagccccg	agacccccaa	ggcagaggga	ngccggcctg	ttggccctcc		540
acgctatccc	tctgcagcct	gggcccctcc	gacagaggcc	ccagggtgcg	tgscagtgra		600
ggtggggcac	ttaggtgcct	ggctggccca	gggcttgctc	tccgtgtcaa	gccgactcac		660
ccagagccca	ccctcccaag	ctcaggggca	tcctccgctg	ggccccagtg	cctttgcrct		720
gcgcagcact	ctgccctcca	ctggactcag	gcattgtctat	ggctgcctgt	cctgaggctc		780
cggagccctc	atttcttctg	gaagtcccca	gctcccctgc	ctccactcaa	tggcaccggc		840
cctgcaactt	taggcaggtc	gaagccaacc	caaggaaaga	acctaagaac	ctcgtttgga		900
gggatgtcag	cttggggccag	mccagccgca	ccccgcgggg	ctcaggcttg	gaactgggtga		960
gggtgtgtgg	tgggggtatg	cagagggata	agaccgtggt	agaggagagg	gttgggtgag		1020
agagagagag	agagagagag	agagagagtc	tggggggagc	gggcaagcat	ggggagatga		1080
gatgtgtata	tgtgagagag	agtgtggggg	ccccaggcag	ggcaggaggt	ggtggaaacg		1140
gggtgaactc	cgtgggctgt	gtgaggactg	tccatagtgg	gtccmaaccc	cctccctctg		1200
ctggagtttc	ctagcccttc	cccctcccya	agactgwggc	agcaggcagg	agccccctgc		1260
ctccctccct	gtcctgtgcc	acacttctgg	ggccaaaccc	agcccccttg	agccaggccc		1320
tgccagactc	caagcccacc	ctagaaccct	cctcctgtgt	ggagactctg	ttgccccact		1380
ttggacacag	attggcaacc	tgccctaccm	ckccccctw	cgctggggct	tccatcttaa		1440
tttattctca	ataataaaga	cttcatgatg	amaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa		1500
aaaaaaaaaa	aa						1512

<210> 19

<211> 1655

<212> DNA

<213> Homo sapiens

<400> 19

ccacgcgtcc	gggcaaagaa	ttaaacctgg	tgtttggact	tcaacttagc	atggctagaa	60
ttggaagtac	agtaaacatg	aacctcatgg	gatggctgta	ttctaagatt	gaagctttgt	120
taggtttctgc	tggtcacaca	accctcggga	tcacacttat	gattgggggt	ataacgtgta	180
ttctttcact	aatctgtgcc	ttggctcttg	cctacttgga	tcagagagca	gagagaatcc	240
ttcataaaga	acaaggaaaa	acaggtgaag	ttattaaatt	aactgatgta	aaggacttct	300
ccttaccctt	gtggcttata	tttatcatct	gtgtctgcta	ttatgttgct	gtgttccctt	360
ttattggact	tgggaaaagt	ttctttacag	agaaatttgg	attttcttcc	caggcagcaa	420
gtgcaattaa	cagtgttgta	tatgtcatat	cagctcccat	gtccccgggt	tttgggtcc	480
tggtggataa	aacagggaag	aacatcatct	gggttctttg	cgcatagcag	ccactcttgt	540
gtccacatg	atgtggcct	ttacgatgtg	gaacccttgg	attgctatgt	gtcttctggg	600
actctcctac	tcattgcttg	cctgtgcatt	gtggccaatg	gtggcatttg	tagttccctga	660
acatcagctg	ggaactgcat	atggcttcat	gcagtccatt	cagaatcttg	ggttggccat	720
catttccatc	attgctggta	tgatactgga	ttctcggggg	tatttgtttt	tggaagtgtt	780
cttcattgcc	tgtgtttctt	tgctactttt	atctgtggtc	ttactctatt	ggggaatcgt	840
gcccagggtg	ggaacctaaa	ttattctgca	agacaaagga	agaaataaaa	tttcccatac	900
tgaatgagaa	gttaaaatga	atgtgtcaga	gaatgggctt	aacacatcgt	tggtttgaaa	960
acttccattt	taaaaattta	gagtttagtc	attagaaaaa	ataatggact	ggaaagttaa	1020
atttatatcc	aaatatacct	atttcaaagt	gtatttgtga	ggcctgtttt	agcctgtgtc	1080
ttttgtattg	tgtgttgcta	aagaattcta	cttttagtag	gctaatacaac	aatgaaaggg	1140
ttagaaaatt	gctgtggaac	atccagggtg	acttcaggaa	agacagtga	aaatggaaaa	1200
cgttggagct	tctgttgaga	taatcttcat	taggtatata	tcttagggat	acagcctttt	1260
ctttatctta	tagcaggaaa	aaaaaacttt	tgagggaat	agaagggtctg	cgttacacaa	1320
aataaacaat	ggcattgtca	taggccttcc	ttttactagt	agggcataat	gctagggaat	1380
atgtgaagat	gtttttttga	agtctcttct	tgatcacgaa	caatagcttg	cgctctactc	1440
tgtagttagt	tggattgccc	agcaatgacc	cttttcaatt	tcttatttct	gtgttactga	1500
ggacccta	cacttaggga	tgtaatttta	tagtataaac	tttctgtaca	gtttttctta	1560
tagtctaata	agtaaaaagt	gtccttcaaa	ttatgataat	tgccatgta	catggataaa	1620
ttaaaacact	gcacacggaa	aaaaaaaaaa	aaaaa			1655

<210> 20

<211> 2525

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (5)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (10)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1354)

<223> n equals a,t,g, or c

<400> 20

tgacnctatn	gtaaggtagc	cctgcaggta	ccgggtccgga	attcccgggt	cgaccacgc	60
gtccgggtctg	ccaacaaggt	cgttcatgaa	agtgtttttc	tctttaaggt	aattaaaaaa	120

cagtggaaatg	gaaaaacagt	gctgtagtca	tccgtgaata	tgctccttgt	caacaatgta	180
tacattcctg	ctaggtgcca	tattcattgc	tttaagctca	agtcgcatct	tactagtga	240
gtattctgcc	aatgaaggta	agttaagact	tggtatatgc	atggagcact	tccatcta	300
cacacatctc	tctcttgctt	ttgggtctgt	tatatataac	atggaaataa	taatgccttt	360
tgcttcatgt	gagtataaaa	gcatatttaa	atttgattat	ttaaccttgc	attcctcaac	420
aagaaaaaat	gtttgataat	ggatgaaatg	tgagtcaatc	agatacaaaa	atcaaacctt	480
ttggtgaaga	accagtcgta	acatttgact	gttaattcaa	tcaacagggtg	ttcttggaac	540
tatagcaaaa	tggtgaattg	cgcttatttt	tgaagtagaa	ggatataattt	gtttggtcac	600
ttggcatttg	tgaggtactt	actattgtaa	ttattgtatc	aatggtaagg	tgtagcattt	660
atattgtgcg	gtcatattgt	atcaacagta	taaattataa	gctttgataa	gtatgtattt	720
aagaaatctt	tttttatgta	gggatttaag	caaacacttt	aattccacca	aactgtattg	780
agtacttctt	actagtattt	gagtgagggtg	gtgggttgcc	cctccacatc	tgtaggtgtt	840
tctcgttagg	tggaacgaga	gacttgga	agaaagggtg	atagacaaag	tatagagaaa	900
gaaaaaagg	ggccagggtg	accggcgctc	agcacacgga	ggatctctgc	cagcctctga	960
gttccmtag	tatttatgta	tcattattgg	gtgtttctcg	gagagtggga	tgtaggcagga	1020
tcataaggata	gtagtggaga	gagggtcaac	aggtaaacac	gtgaacaaag	gtctttgcat	1080
cataracaak	gtaagrattt	aagtgtctgt	cttttagata	tgcatacaca	taaacaatctc	1140
aatgctttac	aaagcagtat	tgctgcccgc	akgtcccacc	tccagcccta	aggcgggtttt	1200
ycctatctc	agtagatgga	gcatacaatc	gggttttata	ccgagacatt	ccattgcccc	1260
gggacrggca	ggagacagat	gccttcctct	tgctcact	gcaagaggcr	ttccttcctc	1320
ttttactaat	cctcctcagc	acagaccctt	tacnggtgtg	cggtctgggtg	gacggtcagg	1380
tctttccctt	cccacgaggtc	catatttcag	actatcacat	ggggagaaac	cttggaacat	1440
acctggtctt	cctaggcaga	ggtcscgtcg	gcyttccrca	gtgttttggtg	tccctgsgta	1500
cttgagatta	gggagtggtg	atgactctta	asgagcatgc	tgcttcaag	catctgttta	1560
acaaagcaca	tcttgaccg	cccttaatcc	atttaaccct	gagtkgacac	agcacatgtt	1620
tcagagagca	crgggttggtg	ggtaagggtta	cagattgcag	aacaaaatgg	agtctcctat	1680
gtctacttct	ttctamacag	acacagtaac	aatctgatct	ctcttttccc	cacaattgag	1740
gacacataca	atcatgatata	gacctttaat	ggtctactac	ttggagagtc	agatgtgtac	1800
ccaagtctct	actgcagtta	acatttacct	gccaggcact	aggctaagta	ttagcagcag	1860
gttcaaagtg	cataagatat	agaccttgct	ctcaagactt	agtttattag	gagagacatg	1920
aatgtaaaaa	catcatgaaa	atccattata	ataactgcaa	taattgatata	atcctgaaga	1980
tgagagagatt	gtctagagga	taaagtata	tattctgttt	ggtaggggat	gatgtggagt	2040
tcaaatggat	cagagaacac	ttcgctgatt	agaagtcagt	tgatccacta	gaagtcaagg	2100
tgaaacaagg	gattcaaaaac	agaggcaaca	gcctgtaaaa	gggaacagag	gcataaaaaa	2160
gcaggatag	ttgtgagaac	atgtagtttg	aaattaccaa	gcaaaaagtt	taaggactcg	2220
tagccaggca	cagtggctca	tgctgtaatc	ccagcacttt	aggaggccaa	ggccggcgga	2280
tcacttgagg	tcaggaattt	gagaccaggtc	tgccaacat	gggaaaccc	atctctacta	2340
gaaatacaaa	aaattagctg	ggtttggttg	cggtgtcctg	taatcccagc	tatgagggag	2400
actgaagcag	gagaatgaac	ccgggaggca	gagattgcag	tgagccgaga	tcattgccact	2460
gcactccagc	ctgggcaaca	gagcaaaact	gtctcaagaa	aaaaaaaaaa	aaaaagggtcg	2520
gccgc						2525

<210> 21

<211> 1396

<212> DNA

<213> Homo sapiens

<400> 21

aagtctcgta	tcgccccggg	gaggcgccgg	agcccagcgg	ctggcgccag	atccagggtc	60
ctggaagaac	catgtccggc	agctactggt	catgccaggc	acacactgct	gcccaagagg	120
agctgctggt	tgaattatct	gtgaatggtg	ggaagaggaa	tgccagagct	gccggctgaa	180
aattacccaa	ccaagagaaa	tctgcaggat	ggactttctg	gtcctcttct	tggtctacct	240
ggcttcgggtg	ctgatgggtc	ttgttcttat	ctgcgtctgc	tcgaaaaccc	atagcttgaa	300
aggcctggcc	aggggaggag	cacagatatt	ttcctgtata	attccagaat	gtcttcagag	360
agccrtgcat	ggattgcttc	attacctttt	ccatacgaga	aaccacacct	tcattgtcct	420
gcacctggtc	ttgcaaggga	tggtttatac	tgagtaacac	tggaaggtat	ttggctactg	480
tcaggagctg	gagttgtcct	tgcatctact	tctctgccc	tatctgctgc	taggtgtaaa	540

10

cctgtttttt	ttcacctga	cttgtggaac	caatcctggc	attataacaa	aagcaaatga	600
attattattt	cttcatgttt	atgaatttga	tgaagtgatg	tttccaaaga	acgtgagggtg	660
ctctacttgt	gattttaaagga	aaccagctcg	atccaagcac	tgcagtgtgt	gtaactgggtg	720
tgtgcaccgt	ttcgaccatc	actgtgtttg	ggtgaacaac	tgcacggggg	cctggaacat	780
caggtaacttc	ctcatctacg	tcttgacctt	gacggcctcg	gctgccaccg	tcgccattgt	840
gagcaccact	tttctggtcc	acttgggtgt	gatgtcagat	ttataaccagg	agacttacat	900
cgatgacctt	ggacacctcc	atgttatgga	cacggctctt	cttattcagt	acctgttcct	960
gacttttcca	cggattgtct	tcatgctggg	ctttgtcgtg	gttctgagct	tcctcctggg	1020
tggttacctg	ttgtttgtcc	tgtatctggc	ggccaccaac	cagactacta	acgagtggta	1080
cagaggtgac	tgggcctggt	gccagcgttg	tccccttgtg	gcctggcctc	cgtcagcaga	1140
gccccaaagt	caccggaaca	ttcactccca	tgggcttcgg	agcaaccttc	aagagatctt	1200
tctacctgcc	tttccatgtc	atgagaggaa	gaaacaagaa	tgacaagtgt	atgactgcct	1260
ttgagctgta	gttcccgttt	atttacacat	gtggatcctc	gttttccaaa	aaaaaaaaaa	1320
aaaaaaaaaa	aaaaaaaaaa	aaaactcgag	ggggggcccg	gtaccaatt	cgccctggag	1380
ttcaagtaga	catcaa					1396

<210> 22

<211> 1069

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (508)

<223> n equals a,t,g, or c

<400> 22

ggcacgagca	cagcctcagg	ccctgcccga	gacctgcaga	atcagaaact	ctgggggtgag	60
gcctggttat	ctgctgtaac	agaccttcca	gtgggttctg	atgccctcta	gagcaggaga	120
accactagct	tagagggtgc	agtatgtttg	gcatcttgcc	atttgtgtta	gttcagagga	180
atggctgacc	cccatgtctc	atttctaagc	ttcaggcagc	ttttctcctg	ggcagctgtc	240
attctgttga	ggggaatcct	ggggactgtg	gctcctcctc	cctgtccgtg	tgtccttgat	300
ctggcagtct	acccccctca	tctccccgtg	gaggctccat	gcctagaggt	ggtcttcaaa	360
cagaagaatg	gcaaagataa	ttgtctcgtg	ttttaccctg	accccatctc	tttaagaggg	420
tcacttcttg	gcccattcat	taaaaaccaa	tgtcatagtt	ctgtgattcc	actatcagac	480
agtgccacgt	ccaaggcgcg	ggctcttnac	ctccctggaa	gagagactgt	gctgtctgtg	540
cttctgtgtg	tctccagtc	cacgtctcca	cggaccacag	cccttgagga	ctccctcggt	600
gtcccagggc	ttctggtgtg	ttcagagacc	tccacactca	acgaccactg	gtgctgcaga	660
agggccgggtg	cttacattcc	aattaacaga	cgtttttccc	atctaattgcc	tcttgccctc	720
tcctaacacc	acctcgggag	tgtttatgtc	tattctaagt	gaatttcact	gtgtgaaaaa	780
attcacacct	gttgctccag	cgatttgga	ggccggggcg	ggtgtatcat	ttgagcccag	840
gagtttgagg	ctagcctggg	caggatgggt	aaaccccgct	tctataaaga	aattttaaaa	900
attagctggg	catagtggca	cgtgcctgta	gttccatcta	ctggggaggc	tggggtgga	960
ggatcgcatg	agcccgagg	tttgaggctg	cagtgaagctg	tgatcgagc	actgcactcc	1020
agtcctggga	acagagcaag	accctgtctc	ttaaaaaaaa	aaaaaaaaaa		1069

<210> 23

<211> 1658

<212> DNA

<213> Homo sapiens

<400> 23

ggcacgagcc	ggcctgccag	agccatgcc	ctgactcctc	agcttcaaaa	tcagggggtct	60
caggacagag	gatgctgggt	gggctcagag	ctcatcaggg	gggctgtgtg	tgagagggga	120
tgccttctgg	atgccctcat	cctcctcggg	gctgggggtc	ccctcaaggc	caccagctc	180
cttcttttgt	ttgtgtgtgc	tactcctgcc	gcctgtctgc	ttggccctgc	tgctcttctt	240

11

cttggacttc	ttccctccca	gggcagctgt	gtctcccttc	ttgcccggacc	actgctctgc	300
caggcaacct	aggggtgtga	ggagagagac	cctcaacaga	agtgcctcag	ggctgggggtg	360
ctgggcaagg	agcactgagc	agggagccgt	gggagtagca	actgggactg	tgcttgacat	420
cagcctccct	gcctcctgcc	tctcgtctgt	gccaccaggc	ccctctgggg	gcactctgact	480
tgtctgcccc	tcattctgca	cctgggtttca	gtgactctta	cttcacccatg	tcttgccaat	540
caagcctttc	aagggcagag	atcctacaat	gccctctggt	gccctctgtt	tctcctccta	600
cccacctccc	ccaagggaga	gcaaacaaat	catccagagc	cagcctgccc	ttgcttcccc	660
aaactcactg	gtgtcttttc	ccttcagcac	gtggttggcg	cagaggaatt	cagtcaggtc	720
ttcctcctgg	tggatcctgt	accagtccct	gatcacctcc	tcaaactctt	caccagcaca	780
tcacacttgt	taatcataat	acctcatatt	ggcaaagccc	cagcacctga	ctcgctccta	840
gaggagctca	gcctaagcct	cgcaaccac	tgcaaggtag	cagtggcacg	gttcacctaa	900
ggaaactgag	gccagagagg	tgaaatgacc	tgaccaaagc	caccccgccc	tggttgact	960
tcctcagagc	agacccaatc	cccaccagcc	cttactggg	cacagcaacc	cttccaaggg	1020
ctgaagggcc	tgtacctgct	tcttgaggtc	agccacttct	gcagaagtct	cgttcaacag	1080
ctcatagggg	atgtccatca	ccaccttgac	ccctttgtgt	accaggttgt	gtaatgtctc	1140
aaaggtctct	gacatgccct	ggaagaagcg	accagatatg	gcaggcggag	ctcccttctc	1200
tcctctccac	cctcgtctcc	cagtgtgtgc	taagaaccca	gctataagac	caatgtctaa	1260
cgccctctaa	ggatcctcat	cctttttttt	ttgagaagga	gtctcactct	gtcgcccagg	1320
ttggagcgtc	tcagctcact	gcaacctctg	cctcccagg	tcaagcgatt	ttcctgcctc	1380
agcctcccaa	gcagctggga	ctacaaagc	gtgccaccat	acccggctaa	ttttgtaga	1440
gttgggtttt	tgtcatgttg	gtcaggctgg	tctcgaaact	ctagcatcaa	gttttccact	1500
cacctcagcc	tcccaaagtg	ctgagattac	aggcgtgagc	caccgcacct	ggcctcatcc	1560
ttgacctgac	cttctctctc	cctcttttag	gcctgcttcc	cacaaccctt	gcacatatat	1620
ccctgatct	gcctctgcac	acctcatcgc	ttcaaaaa			1658

<210> 24

<211> 1077

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1036)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1038)

<223> n equals a,t,g, or c

<400> 24

ggcacgaggg	gaaagccatg	ctcccaggac	tccttccttg	cagccttaaa	tcggtctgta	60
cggaaaattc	cgcgcttag	aaaccacgc	ttgggtgtaa	cttattattg	ttcttctga	120
cctacttcct	gtttatcact	tccgggttca	tcattttggc	atttcggtga	tcgggttgga	180
actattgaag	cccgctttca	ggttcttttc	cccattttcc	ctttgaaagg	aagacttctg	240
gcttctccta	aatctccgtt	ctctgggtaa	ggggagtcca	agcctctgtc	atgaggaacg	300
gaaatgcgag	ggcctcgggt	gttactctaa	aatccgccct	cagcttgac	gccggaagct	360
gcgattcctg	cagcgggaaga	ggcgtgatct	ggccttcgac	tcgctatgtc	cactaacaat	420
atgtcggacc	cacggaggcc	gaacaaagtg	ctgaggtaga	agccccgcc	gagcgaatgt	480
aacccggcct	tggacgaccc	gacgccggac	tacatgaacc	tgctgggcat	gatcttcagc	540
atgtgcggcc	tcattgcttaa	gctgaagtgg	tgtgcttggg	tcgctgtcta	ctgctccttc	600
atcagctttg	ccaactctcg	gagctcggag	gacacgaagc	aaatgatgag	tagcttcatg	660
ctgtccatct	ctgccgtggt	gatgtcctat	ctgcagaatc	ctcagcccat	gacgccccca	720
tggtgatacc	agcctagaag	ggtcacattt	tggaacctgt	ctatccacta	ggcctgggct	780
ttggctgcta	aacctgctgc	cttcagctgc	catcctggac	ttccctgaat	gaggccgtct	840
cggtgcccc	agctggatag	agggaaacctg	gcccttccct	agggaaacacc	ctaggcttac	900
ccctcctgcc	tccttctccc	tgctgtctgc	tgggggagat	gctgtccatg	tttctagggg	960

12

tattcatttg	ctttctcgtt	gaaacctgtt	gttaataaag	tttttcactc	tgaaaaaaa	1020
aaaaaaaaa	aaaaancncg	agggggggcc	cggaacccaa	ttccscggat	agtgagt	1077

<210> 25
 <211> 1205
 <212> DNA
 <213> Homo sapiens

<400> 25

cccacgcgtc	cgcagaggca	gggcaatagt	ggagtctctg	cttggccaag	cagcctagaa	60
ctcaaagtcc	atggcccctt	ctgggccttg	agaaattgga	tggttatagc	accaggcagc	120
ccttgtgggt	gggggacagc	aaatgaggga	cctctctttt	ctctacactc	tcctttggct	180
cccggagatc	tggcaggccc	tggctggagg	cataagatta	gatgagggtg	agctgttgga	240
gaatgaagct	gtgttgggag	aagaaatgag	gttgtaccgg	aagatcaacg	agggtgtgct	300
gtcagggaat	gagggtgtac	ttgggggcaa	gtgaggctgc	attattagat	aaatgagggt	360
gtactgtcag	gggatgaagt	gtactttag	tagagatgac	gtcctgctgg	atcagtcggc	420
ttttgtcca	tcagagaaca	cagccacacc	acaggaggaa	ggagagtgtc	cgactcagag	480
gataaatgag	ggtgtcctgc	tggataaatg	agggggcccg	tcaggatgaat	ggagtgtctg	540
tagcaaatga	ggtgtacttt	gctggataaa	tgggactggg	gtgctggata	aatgggggtg	600
tgctgtcagg	tgaatgcatt	actgtctctg	ggtgaagggc	atcctgggaa	taatgagggt	660
gtcctgtctg	atagatgagc	tgccaccacc	aaatggatca	gacctgtctc	atgaaggagg	720
caccatcagc	aacgacgagg	ttatcctgtt	cccactgggg	ctcctggagc	gtcttctggc	780
ccaggggaaa	ctcgggtgtg	gccaccctgg	gttatccaag	tctctctggg	gagcagggtg	840
gggggctggg	gagggcaggc	agctgcattg	tgcaccgtgg	gacctctcct	tcacccccaa	900
tggatgccct	actcctctcc	ctggcacccc	tcagtgggtc	agactgcttc	ggacattctc	960
acccactgc	ctgcttctca	tcttgcctgt	gtcttctttc	tgcccagttt	ggaaaagccc	1020
ctattatgtg	tcagccactc	tgccagtcct	tatttaattct	ccctataaca	cagtattact	1080
cctccttgca	catacacact	ttctcttatt	cattcatcca	ttcattcatt	tgacaaacat	1140
ttaagtgtct	agtatgtacc	aaacacatga	ggtacagttt	taaaaagaat	aaaaaaaaaa	1200
aaaaa						1205

<210> 26
 <211> 1674
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (1663)
 <223> n equals a,t,g, or c

<400> 26

cccagagcagc	tgagtccctt	ccctgtcttt	cactcttctg	gcatcgggtg	ttttacttct	60
tcgattgaac	cctgtcttct	cgacccccct	gggaggccgc	cttcttcagg	cgcctccctt	120
ctctccacga	gctcgtctct	acagctgagg	aactggcaag	atcctgctac	ccagagggtg	180
aatgggtatc	tttcccggaa	taatcctaata	ttttctaagg	gtgaagtgtg	caacggcggc	240
cgtgattgta	agcggagtaa	gcaaacacct	ccattgtatt	agtcaccaga	aaagtaccac	300
tgtaagtcat	gagatgtctg	gtctgaattg	gaaacccttt	gtatatggcg	gccttgccctc	360
tatcgtggct	gagtttggga	ctttccctgt	ggaccttacc	aaaacacgac	ttcaggttca	420
aggccaaaagc	attgatgccc	gtttcaaaaga	gataaaatat	agagggatgt	tccatgcgct	480
gtttcgcctc	tgtaaagagg	aagggtgtatt	ggctctctat	tcaggaattg	ctcctgcggt	540
gctaagacaa	gcatcatatg	gcaccattaa	aattgggatt	taccaaagct	tgaagcgctt	600
attcgtagaa	cgtttagaag	atgaaactct	tttaattaat	atgatctgtg	gggtagtgtc	660
aggagtgata	tcttccacta	tagccaatcc	caccgatgtt	ctaaagattc	gaatgcaggc	720
tcaagggaagc	ttgttccaag	ggagcatgat	tggaaagcttt	atcgatatat	accaacaaga	780
aggcaccagg	ggtctgtgga	gggggtgtgt	tccaactgct	cagcgtgctg	ccatcgttgt	840

13

```

aggagtagag ctaccagtct atgatattac taagaagcat ttaatattgt caggaatgat      900
gggcgataca attttaactc acttcgtttc cagctttaca tgtggtttgg ctggggctct      960
ggcctccaac ccggttgatg tgggtcgaac tcgcatgatg aaccagaggg caatcgtggg    1020
acatgtggat ctctataagg gcactgttga tggatattta aagatgtgga aacatgaggg    1080
cttttttgca ctctataaag gattttggcc aaactggcct cggcttggaac cctggaacat    1140
catttttttt attacatacg agcagctaaa gaggtctcaa atctaagaac tgaattatat    1200
gtgagcccgag cccctgccagc ctttctactc ctttgccttt ttcccgtgtt ctaatgtatt    1260
ttgacaatgt tgtaagtgtt taccaagccg ttgggtctcct aagggcctcc tgatggaaga    1320
acagtggggg ggttcaaagt tatttctatg tttgtgttac catgttaact tttcccggag    1380
agaagtgtt aacattgaga ctctggcccc agattgggat cttctatgaa gatggatact    1440
gatgggtgac attgaaaacg gcctgctttc caaatgtggt taaatgtaat tgggttagccc    1500
cagacttggg ctagagcaga aggcataagg caggggtggt attgctatat gtgttacaga    1560
cctcggttct cattaagta tttattggca gaatcaaaaa aaaaaaaaaa aaaaaaaaaa    1620
aaactcgagg gggggcccgg tacccaattc gccctatggt gantcgaatg ggct      1674

```

<210> 27
 <211> 1965
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (333)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1961)
 <223> n equals a,t,g, or c

```

<400> 27
ggatcctcgc ggcggcggcg gtgcttacag cctgagaaga gcgtctcgcc cgggagcggc      60
ggcggccatc gagaccacc caaggcgcgt cccctcgcgc ctccagcgc tcccaagccg      120
cagcggccgc gcccttcag ctagctcgcct cgctcgcctc gcttcctcgc tgcggcgtgc      180
gcatggcggt ggcgttgggc gcgctggcgg cggctcgagcc ggctgcaggc agccgggtacc      240
agcagttgca gaatgaagaa gactctggag aacctgaaca ggctgcaggc gatgctcctc      300
caccttacag cagcatttct gcagagagcg cancatattt tgactacaag gatgagtctg      360
ggtttccaaa gccccatctt tacaatgtag ctacaacact gcccagttat gatgaagcgg      420
agaggaccaaa ggctgaagct actatccctt tgggtcctgg gagagatgag gattttggg      480
gtcgggatga ttttgatgat gctgaccagc tgaggatagg aaatgatggg attttcatgt      540
taactttttt catggcattc ctctttaact ggattgggtt tttcctgtct ttttgctga      600
ccacttcagc tgcaggaagg tatggggcca tttcaggatt tgggtctctc ctaattaaat      660
ggatcctgat tgtcagggtt tccacctatt tccctggata ttttgatggt cagtactggc      720
tctggtgggt gttccttggt ttaggccttc tccgtttctc cagaggattt atcaattatg      780
caaaagttcg gaagatgccg gaaactttct caaatctccc caggaccaga gttctcttta      840
tttattaaag atgttttctg gcaaaggcct tccctgcattt atgaattctc tctcaagaag      900
caagagaaca cctgcaggaa gtgaatcaag atgcagaaca cagaggaata atcacctgct      960
ttaaaaaaat aaagtactgt tgaaaagatc atttctctct atttgctcct aggtgtaaaa    1020
ttttaatagt taatgcagaa ttctgtaatc attgaatcat tagtggttaa tgtttgaaaa    1080
agctcttgca atcaagtctg tgatgtatta ataatgcctt atattattgt ttagtcatt      1140
ttaagtagca tgagccatgt ccctgtatgc ggtagggggc agtcttgctt tattcatcct      1200
ccatctcaaa atgaacttgg aattaaatat tgtaagatat gtataatgct ggccatttta      1260
aagggttttt ctcaaaagtt aaacttttgt tatgactgtg tttttgcaca taatccatat      1320
ttgctgttca agttaatcta gaaatttatt caattctgta tgaacacctg gaagcaaaat      1380
catagtgcaa aaatacattt aaggtgtggt caaaaaaag tctttaattg gtaataata      1440
agcattaatt ttttatagcc tgtattcaca attctcggtt accttattgt acctaaggga      1500
ttctaaaggt gttgtcactg tataaaacag aaagcactag gatacaaatg aagcttaatt      1560

```

14

actaaaaatgt	aattcttgac	actctttcta	taattagcgt	tcttcacccc	cacccccacc	1620
ccccccccc	ttattttcct	tttgtctcct	ggtgattagg	ccaaagtctg	ggagtaagga	1680
gaggattagg	tacttaggag	caaagaaaga	agtagcttgg	aacttttgag	atgatcccta	1740
acatactgta	ctacttgctt	ttacaatgtg	ttagcagaaa	ccagtgggtt	ataatgtaga	1800
atgatgtgct	ttctgcccga	gtggtaattc	atcttgggtt	gctatgttaa	aactgtaaat	1860
acaacagAAC	attaataaat	atctcttggt	tagcaccttt	aaaaaaaaa	aaaaaaaaa	1920
aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	naaaa		1965

<210> 28

<211> 1863

<212> DNA

<213> Homo sapiens

<400> 28

gactaggcgc	cgagcttagt	cctggggagcc	gcctccgtcg	ccgccgtcag	agccgcctca	60
tcagattatc	ttaacaagaa	aaccaactgg	aaaaaaaaa	gaaattcctt	atcttcgcat	120
ttttcgggtg	tggtcacctt	ttatccctgt	gctctgggaa	agctatatgc	aagaatggca	180
tctctaagag	gacttttgaa	gaaataaaag	aagaaatagc	cagctgtgga	gatgttgcta	240
aagcaatcat	caacctagct	gtttatggta	aagcccagaa	cagatcctat	gagcgattgg	300
cacttctggt	tgatactggt	ggacccagac	tgagtggctc	caagaacctt	gaaaaagcca	360
tccaaattat	gtaccaaAAC	ctgcagcaag	atgggctgga	gaaagtccac	ctggagccag	420
tgagaatacc	ccactgggag	aggggagaag	aatcagctgt	gatgctggag	ccaagaattc	480
ataagatagc	catcctgggt	cttggcagca	gcattgggac	tcctccagaa	ggcattacag	540
cagaagtctt	ggtggtgacc	tctttcgatg	aactgcagag	aagggcctca	gaagcaagag	600
ggaaagattg	tgtttataac	caaccttaca	tcaactactc	aaggacggtg	caataccgaa	660
cgcagggggc	ggtggaagct	gccaaggttg	gggctttggc	atctctcatt	cgatccgtgg	720
cctccttctc	catctacagt	cctcacacag	gtattcagga	ataccaggat	ggcgtgcccc	780
agattccaac	agcctgtatt	acggtggaag	atgcagaaat	gatgtcaaga	atggcttctc	840
atgggatcaa	aattgtcatt	cagctaaaga	tgggggcaaa	gacctacca	gatactgatt	900
ccttcaacac	tgtagcagag	atcactggga	gcaaatatcc	agaacagggt	gtactggtca	960
gtggacatct	ggacagctgg	gatgttgggc	agggtgccat	ggatgatggc	ggtggagcct	1020
ttatatcatg	ggaagcactc	tcacttatta	aagatccttg	gctgcgtcca	aagaggactc	1080
tgcggtggtg	gctctggact	gcagaagaac	aaggtggagt	tggtgccttc	cagtattatc	1140
agttacacaa	ggtaaatatt	tccaactaca	gtctggtgat	ggagtctgac	gcaggaaacct	1200
tcttaccac	tgggctgcaa	ttcactggca	gtgaaaaggc	cagggccatc	atggaggagg	1260
ttatgagcct	gctgcagccc	ctcaatatca	ctcaggctct	gagccatgga	gaagggacag	1320
acatcaactt	ttggatccaa	gctggagtgc	ctggagccag	tctacttgat	gacttataca	1380
agtattttct	cttccatcac	tcccacggag	acaccatgac	tgatcatgat	ccaaagcaga	1440
tgaatgttgc	tgctgctggt	tgggctgttg	tttcttatgt	tggtgcagac	atggaagaaa	1500
tgctgcctag	gtcctagaaa	cagtaagaaa	gaaacgtttt	catgcttctg	gccaggaatc	1560
ctgggtctgc	aactttggaa	aactcctctt	cacataacaa	tttcatccaa	ttcatcttca	1620
aagcacaaact	ctatttcagt	ctttctgtta	ttatctttct	tgatactttc	caaattctct	1680
gattctagaa	aaaggaatca	ttctcccctc	cctcccacca	catagaatca	acatatggta	1740
gggattacag	tgggggcatt	tctttatata	acctcttaaa	aacattgttt	ccactttaaa	1800
agtaaacact	taataaatat	ttggaagatc	tctgaaaaaa	aaaaaaaaa	aaagggcggc	1860
cgc						1863

<210> 29

<211> 1626

<212> DNA

<213> Homo sapiens

<400> 29

cccacgcgtc	cggagccggg	agccgggtcg	gggggctccg	ggctgtggga	ccgctggggc	60
cccagcgatg	gcgaccctgt	ggggaggcct	tcttcggctt	ggctccttgc	tcagcctgtc	120
gtgcctggcg	ctttccgtgc	tgctgctggc	gcactgtcag	acggcgccaa	gtgattgcct	180

15

tcatgtttgtg	gagcccatgc	ctgtgctggg	gcctgatgta	gaagcatact	gtctacgctg	240
tgaatgcaaa	tatgaagaaa	gaagctctgt	cacaatcaag	gttaccatta	taatttatct	300
ctccattttg	ggcctttctac	ttctgtacat	ggtatatctt	actctgggtg	agcccatact	360
gaagaggcgc	ctctttggac	atgcacagtt	gatacagagt	gatgatgata	ttggggatca	420
ccagcctttt	gcaaatgcac	acgatgtgct	agcccgcctc	cgagctcgag	ccaacgtgct	480
gaacaaggta	gaatatgcac	agcagcgctg	gaagcttcaa	gtccaagagc	agcgaaagtc	540
tgtctttgac	cggcatgttg	tcctcagcta	attgggaatt	gaattcaagg	tgactagaaa	600
gaaacaggca	gacaactgga	aagaactgac	tgggttttgc	tgggtttcat	tttaatacct	660
tgttgatttc	accaactggt	gctggaaatt	caaaactgga	agcaaaaact	tgcttgattt	720
ttttttcttg	ttaacgtaat	aatagagaca	tttttaaaag	cacacagctc	aaatcagcca	780
ataatctttt	cctattgtga	cttttactaa	taaaaataaa	tctgcctgta	aattatcttg	840
aagtccttta	cctggaacaa	gcactctctt	tttcaccaca	tagttttaac	ttgactttca	900
agataatttt	cagggttttt	gttgttggtg	ttttttgttt	gtttgttttg	gtgggagagg	960
ggagggatgc	ctgggaagtg	gttaacaact	tttttcaagt	cactttacta	aacaaaactt	1020
tgtaaataga	ccttaccttc	tattttcgag	tttcatttat	attttgcagt	gtagccagcc	1080
tcatcaaaga	gctgacttac	tcatttgact	tttgactga	ctgtattatc	tgggtatctg	1140
ctgtgtctgc	actttcatgt	aaacgggac	taaaatgcct	gggtggcttt	cacaaaaagc	1200
agattttctt	catgtactgt	gatgtctgat	gcaatgcac	ctagaacaaa	ctggccattt	1260
gctagtttac	tctaaagact	aaacatagtc	ttggtgtgtg	tgggtcttact	catcttctag	1320
tacctttaag	gacaaatcct	aaggacttgg	acacttgcaa	taaagaaatt	ttattttaaa	1380
cccaagcctc	cctggattga	taatatatac	acatttgtca	gcatttccgg	tcgtggtgag	1440
aggcagctgt	ttgagctcca	atgtgtgcag	ctttgaacta	gggctggggt	tgtgggtgcc	1500
tcttctgaaa	ggtctaacca	ttattggata	actggctttt	ttcttctctt	ttggaatgta	1560
acaataaaaa	taatttttga	aacatcaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1620
aaaaaa						1626

<210> 30

<211> 605

<212> DNA

<213> Homo sapiens

<400> 30

ccacgcgtcc	gcccacgcgt	ccgggaaatg	accttggaga	ttgtagcaga	gagtgagcat	60
gaggagcggc	ctgctggcca	gggcccggat	gagcccaaca	tgaaccctaa	gcttgaggac	120
ccaaggcgcc	ccgacacctc	cttctctgtg	tttacctccc	catacaagac	catgaagttc	180
atcctgtggc	ggcgtttccg	gtgggccatc	atcctcttca	tcactctctt	catcctgctg	240
ctgttctctg	ccatcttcat	ctacgccttc	ccgaactatg	ctgccatgaa	gctgggtgaag	300
cccttcagct	gaggactctc	ctgcccctga	gaaggggccc	tgggggtccc	tccagcatgg	360
gactggcctg	cctcctccgc	ccagctcggc	gagctcctcc	agacctccta	ggcctgattg	420
tcctgccagg	gtgggacagc	agacagatgg	accggccccc	actcccagag	ttgctaacat	480
ggagctctga	gatcacccca	cttccatcat	ttccttctcc	cccaacccaa	cgcttttttg	540
gatcagctca	gacatatttc	agtataaaac	agttggaacc	acaaaaaaa	aaaaaaaaaa	600
aaaaa						605

<210> 31

<211> 931

<212> DNA

<213> Homo sapiens

<400> 31

gagagtgcct	aagcgggggt	gaaagaggac	gtgttaccca	ctgccatgca	ccaggactgg	60
ctgtgtaacc	ttgggtggcc	cctgctgtct	ctctgggctg	cagagtctgc	cccacatgtg	120
gccatggcct	ctgcaactgc	tcagctcttg	tccaggccct	gtggcaggac	acacatggtg	180
agcctagccc	tgggacatca	ggagactggg	ctctggctct	gttcggcctt	tgggtgtgtg	240
gtggattctc	cctgggcctc	agtgtgcca	tctgtaaagg	ggcagctgac	agtttgtggc	300
atcttgccaa	gggtccctgt	gtgtgtgtat	gtgtgtgcat	gtgtgcgtgt	ctccatgtgc	360

16

```

gtccatattt aacatgtaaa aatgtccsc ccrckegtcg cccaaacatg ttgtacattt 420
caccatggcc cctcatcat agcaataaca ttcccactgc caggggttct tgagccagcc 480
aggccctgcc agtggggaag gaggccaagc agtgccctgc tatgaaattt caacttttcc 540
tttcatacgt ctttattacc caagtcttct cccgtccatt ccagtcaaat ctgggctcac 600
tcacccacgc gagctctcaa atccctctcc aactgcctaa ggccctttgt gtaagggtgc 660
ttaatactgt cctttttttt tttttaacag tgtttttag atttcagatg actatgcaga 720
ggcctggggg acccctggct ctgggcccgg cctggggctc cgaaattcca aggccagac 780
ttgcgggggg tgggggggta tccagaattg gttgtaata ctttgcatat tgtctgatta 840
aacacaaaca gacctcaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 900
aaaaaaaaa aaaaaaaaaa aaaggcgccg c 931

```

<210> 32

<211> 1407

<212> DNA

<213> Homo sapiens

<400> 32

```

gaattcggca cgagggcagg ctcagaagac gatgcggggc tgtgtgccgg ccttcttgct 60
gcatgtactc agcctcagga gagcttgctg caccagggcc gccaggtct tcacagcaca 120
actgcctgga aggcagggtg cgagaaggag aggcggatgg catgagcagc aagggggacc 180
gatgctgtgc agctcacacc actccagaac ctgacaaggc accagcagga ccccttgcca 240
ggagcatgtc tgtgcagcag tgtttttgcc cctgcacatt ccagaagccc tcatgggaag 300
ggatgcagcc aggcagactc ctgccagatg gggcaggtag tttattcaaa gagaactctg 360
tatcccatag gccagggctc tcctttcgct tggcggtggc tttgctggcc cagtgtgtgc 420
tcctggctca gcagaaacat ccatttgagt tggcatccct gtagggatcc cagagcgttg 480
taagccttct tgtgattggt agggatggct gtgggtggc ttccaggagg gggccaccat 540
tgccgcatct acttctagac tcccaaagga gccagggctc aggcaggcct ggcccagagt 600
cacgctggca accacagagt tgggaagcag tctgtattct tctctctctc tctctctctc 660
agtatccatg acaggatga aacatattgt ctctttataa atgtcatttt acaaaattatg 720
tgattatctg gaagctctaa gatgagagca aatgcctgat cactctggcc aaatgtcaga 780
tactaaagcc cattcttggc cgggcatgtt ggctcccgcc kgtaatccca gcactttggg 840
aagcccaagt gggatgaatc cctgaggtca ggagttcaag accagcctga ccaacatggg 900
gatacccgct ctctactaaa aatacaagcc gggcgtgggt gcgcatgcct gtaatcccag 960
ctactcagga ggctgaggca ggaaatcac ttgaactcgg gaggcagagg ttgcagttag 1020
ctgagatcgc gccattgcac tccagcctgg gtgacagagc aagactctgt ctcataaata 1080
aatacaaaag ccattcttcc agagtcttgt gccttaataa aaacacacct ctctgctgtg 1140
ggaagactgt gcaatggcac agccgcagag cttggtttgg gaggttgaag tgctctgggg 1200
agaattcgta gatcatctc agaaaagcct tgccctgggt ttctaccaga aaaacgtctc 1260
ccaatcacc aggaagctg tccacagtag tccccctta tccacgggtg cactttccat 1320
gggttcagtt atctgcggtc aaccacggtc tgacaatatt aaatggaaaa ttcttcaaac 1380
agttaaaaaa aaaaaaaaaa aactcga 1407

```

<210> 33

<211> 1526

<212> DNA

<213> Homo sapiens

<400> 33

```

ggcacgagaa aaaaccttca ggcgggcccat gggatatgcc aagaggacag agaacgaatg 60
cacagaaata ttgtcagcct tgcacagaat ctccctgaact ttatgattgg ctctatcttg 120
gatttatggc aatgcttctc ctgggttttac attggttctt cattgaatgg tactcgggga 180
aaaagagttc cagcgcactt ttccaacaca tcactgcatt atttgaatgc agcatggcag 240
ctattatcac cttacttttg agtgatccag ttgggtttct ttatattcgt tcatgtcgag 300
tattgatgct ttctgactgg tacacgatgc tttacaaccc aagtcagat tacgttacca 360
cagtacactg tactcatgaa gccgtctacc cactatatac cattgtattt atctattacg 420
cattctgctt ggtattaatg atgctgctcc gacctcttct ggtgaagaag attgcatgtg 480

```

17

ggttaggga	atctgatcga	tttaaaagta	tttatgctgc	actttacttc	ttcccaattt	540
taaccgtgct	tcaggcagtt	ggtggaggcc	ttttatatta	cgcttccca	tacattatat	600
tagtggtatc	tttggttact	ctggctgtgt	acatgtctgc	ttctgaaata	gagaactgct	660
atgatcttct	ggtcagaaag	aaaagactta	ttgttctctt	cagccactgg	ttacttcatg	720
cctatggaat	aatctccatt	tccagagtgg	ataaacttga	gcaagatttg	cccccttgg	780
ctttggtacc	tacaccagcc	cttttttact	tgttcactgc	aaaatttacc	gaaccttcaa	840
ggatactctc	agaaggagcc	aatggacact	gagtgtagac	atgtgaaatg	ccaaaaacct	900
gagaagtgct	cctaataaaa	aagtaaatca	atcttaacag	tgtatgagaa	ctattctatc	960
atatatggga	acaagattgt	cagtatatct	taatgtttgg	gtttgtcttt	gtttgtttta	1020
tggttagact	tacagacttg	gaaaatgcaa	aactctgtaa	tactctgtta	cacagggtaa	1080
tattatctgc	tacactggaa	ggcgcctagg	aagcccttgc	ttctctcaac	agttcagctg	1140
ttcttttagg	caaaatcatg	tttctgtgta	cctagcaatg	tgttccatt	ttattaagaa	1200
aagctttaac	acgtgtaatc	tgcatctctt	aacagtggcg	taattgtacg	tacctgttgt	1260
gtttcagttt	gtttttcacc	tataatgaat	tgtaaaaaa	aacataactg	tgggggtctga	1320
tagcaaacat	agaaatgatg	tatatgtttt	tttgttatct	atctattttc	atcaatacag	1380
tattttgatg	tattgcaaaa	atagataata	atttatataa	caggttttct	gtttatagat	1440
tggttcaaga	tttgtttgga	ttattgttcc	tgtaaagaaa	acaataataa	aaagcttacc	1500
tacataaaaa	aaaaaaaaaa	aaaaaa				1526

<210> 34
 <211> 1737
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (1674)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1731)
 <223> n equals a,t,g, or c

<400> 34						
gtcgaccac	gcgtccgccc	acgcgtccgc	ccacgcgtcc	ggttttataaa	cagaagttaa	60
aaattgtaag	cttaagcttc	cgtttataaa	cagaagttaa	aaattatagg	tcctgttttaa	120
caattcagctc	tgtaaactca	ctcatctttt	tggtttttta	cactttgtca	agatttcttt	180
acataattcat	caatgtctga	agaagttact	tatgcagatc	ttcaattcca	gaactccagt	240
gagatggaaa	aaatcccaga	aattggcaaa	tttggggaaa	aagcacctcc	agctccctct	300
catgtatggc	gtccagcagc	cttgtttctg	actcttctgt	gccttctgtt	gctcattgga	360
ttgggagtct	tggcaagcat	gtttcatgta	actttgaaga	tagaaatgaa	aaaaatgaac	420
aaactacaaa	acatcagtga	agagctccag	agaaatattt	ctctacaact	gatgagtaac	480
atgaatatct	ccaacaagat	caggaacctc	tccaccacac	tgcaaaacaat	agccaccaaa	540
ttatgtcgtg	agctatatag	caaagaacaa	gagcacaat	gtaagccttg	tccaaggaga	600
tggatttggc	ataaggacag	ctgttatttc	ctaagtgatg	atgtccaaac	atggcaggag	660
agtaaaatgg	cctgtgctgc	tcagaatgcc	agcctgttga	agataaaca	caaaaatgca	720
ttggaattta	taaaatccca	gagtagatca	tatgactatt	ggctgggatt	atctcctgaa	780
gaagattcca	ctcgtggtat	gagagtggat	aatataatca	actcctctgc	ctgggttata	840
agaaacgcac	ctgacttaaa	taacatgtat	tgtggatata	taaataagact	atatgttcaa	900
tattatcact	gcacttataa	acaaagaatg	atatgtgaga	agatggccaa	tccagtgcag	960
cttggttcta	catattttag	ggaggcatga	ggcatcaatc	aaatacattg	aaggagtgtg	1020
kggggtgggg	gttctaggct	ataggtaaat	ttaaatattt	tctgggtgac	aattagttga	1080
gtttgtctga	agacctggga	ttttatcatg	cagatgaaac	atccaggtag	caagcttcag	1140
agagaataga	ctgtgaatgt	taatgccaga	gaggtataat	gaagcatgtc	ccacctccca	1200
ctttccatca	tggcctgaac	cctggaggaa	gaggaagtcc	atccagatag	tgtggggggc	1260
cttcgaattt	tcattttcat	ttacgttctt	ccccttcttg	ccaagatttg	ccagaggcaa	1320

18

catcaaaaac	cagcaaattt	taattttgtc	ccacagcggt	gctaggggtg	catggctccc	1380
catctcgggt	ccatcctata	cttccatggg	actccctatg	gctgaaggcc	ttatgagtca	1440
aaggacttat	agccaattga	ttgtttcagg	ccaggtaaga	atggatatgg	acatgcattt	1500
attacctctt	aaaattatta	ttttaagtaa	aagccaataa	acaaaaacga	aaaggcaagt	1560
tacgagactg	acttattttt	aacttctgtg	tgttgagcta	ctgtaagctt	ggcttttgtt	1620
aaagacatac	agcaatttag	tatgcaaaac	taagcattgt	tctgaaaaaa	aatntataga	1680
tagatatgtt	tatctcccat	aactcataac	tggggagtat	tatacccccg	nggcttt	1737

<210> 35

<211> 2242

<212> DNA

<213> Homo sapiens

<400> 35

tcgacccacg	cgtccgggct	gccatggcgg	cggcgggcgg	gctcccagac	tcctggggccc	60
tcttctcgcc	gctcctcgca	gggcttgcac	tactgggagt	cgggcccgtc	ccagcgcggg	120
cgctgcacaa	cgtcacggcc	gagctctttg	gggcccaggg	ctggggcacc	cttgccgctt	180
tcggggacct	caactccgac	aagcagacgg	atctcttcgt	gctgcgggaa	agaaatgact	240
taatcgtctt	tttgccagac	cagaatgcac	cctattttaa	acccaaagta	aagggtatctt	300
tcaagaatca	cagtgcattg	ataacaagtg	tagtccctgg	ggattatgat	ggagattctc	360
aaatggatgt	ccttctgaca	tatcttccca	aaaattatgc	caagagtga	ttaggagctg	420
ttatcttctg	gggacaaaat	caaacattag	atcctaacaa	tatgaccata	ctcaatagga	480
cttttcaaga	tgagccacta	attatggatt	tcaatggtga	tctaattcct	gatatttttg	540
gtatcacaaa	tgaatccaac	cagccacaga	tactattagg	agggaaattha	tcatggcatc	600
cagcattgac	cactacaagt	aaaatgcgaa	ttccacattc	tcatgcattt	attgatctga	660
ctgaagattt	tacagcagat	ttattcctga	cgacattgaa	tgccaccact	agtaccttcc	720
agtttgaaat	atgggaaaat	ttggatggaa	acttytstgw	magtacymta	ttggaaaaac	780
ctcaaaatat	gatgggtggt	ggacagtcag	catttgcaga	ctttgatgga	gatggacaca	840
tggatcattt	actgccaggc	tgtgaagata	aaaattgcc	aaagagtacc	atctacttag	900
tgagatctgg	gatgaagcag	tgggttccag	tcctacaaga	tttcagcaat	aagggcacac	960
tctggggctt	tgtgccattt	gtggatgaac	agcaaccaac	tgaaataacca	attccaatta	1020
cccttcatat	tgagacttac	aatatggatg	gctatccaga	cgctctggtc	ataactaaaga	1080
acacatctgg	aagcaaccag	caggcctttt	tactggagaa	cgtcccttgt	aataatgcaa	1140
gctgtgaaga	ggcgcgctga	atgtttaaag	tctactggga	gctgacagac	ctaaatcaaa	1200
ttaaggatgc	catgggttgc	accttctttg	acatttacga	agatggaatc	ttggacattg	1260
tagtgctaag	taaaggatat	acaaagaatg	atthttgcat	tcatacacta	aaaaataact	1320
ttgaagcaga	tgcttatttt	gttaaaagtta	ttgttcttag	tgggtctgtg	tctaattgact	1380
gtcctcgtaa	gataacaccc	tttggagtga	atcaacctgg	accttatatc	atgtatacaa	1440
ctgtagatgc	aaatgggtat	ctgaaaaatg	gatcagctgg	ccaactcagc	caatccgcac	1500
atttagctct	ccaactacca	tacaacgtgc	ttggtttagg	tcggagcgca	aattttcttg	1560
accatctcta	cgttgggtatt	ccccgtccat	ctggagaaaa	atctatacga	aaacaagagt	1620
ggactgcaat	cattccaaat	tcccagctaa	ttgtcattcc	ataccctcac	aatgtccctc	1680
gaagtgggag	tgccaaaactg	tatcttacac	caagtaatat	tgttctgctt	actgctatag	1740
ctctcatcgg	tgtctgtgtt	ttcatcttgg	caataattgg	cattttacat	tggcaggaaa	1800
agaaagcaga	tgatagagaa	aaacgacaag	aagcccaccg	gtttcatttt	gatgctatgt	1860
gacttgccct	taatattaca	taatggaatg	gctgttcact	tgattagtgt	aaacacaaat	1920
tctggcttga	aaaaataggg	gagattaaat	attattttata	aatgatgtat	cccatggtaa	1980
ttattggaaa	gtattcaaat	aaatatgggt	tgaatatgtc	acaaggctct	tttttttaaa	2040
gcactttgta	tataaaaaat	tgggttctct	attctgtagt	gctgtacatt	ttgttcctt	2100
tgtggaatgt	gttgcatgta	ctccagtgtt	tgtgtattta	taatcttatt	tgcatcatga	2160
tgatggaaaa	agttgtgtaa	ataaaaaata	ttaaatgagc	aggaaaaaaa	aaaaaaaaaa	2220
aaaaaaaaaa	aagggcggcc	gc				2242

<210> 36

<211> 2235

<212> DNA

<213> Homo sapiens

<400> 36

gtaattcggc	acgaggggttc	caccaacatg	gagctctcgc	agatgtcgsa	gctcatgggg	60
ctgtcgggtgt	tgcttgggct	gctggccctg	atggcgacgg	cgccgggtasc	gcgggggtgg	120
ctgcgcgcgg	gggaggagag	gagcggcccg	cccgcctgcc	aaaaagcaaa	tggatttcca	180
cctgacaaat	cttcgggatc	caagaagcag	aaacaatatc	agcggattcg	gaaggagaag	240
cctcaacaac	acaacttcac	ccaccgcctc	ctggctgcag	ctctgaagag	ccacagcggg	300
aacatatctt	gcatggactt	tagcagcaat	ggcaaatacc	tggctacctg	tgcatgatgat	360
cgcaccatcc	gcatctggag	caccaaggac	ttcctgcagc	gagagcaccg	cagcatgaga	420
gccaacgtgg	agctggacca	cgccaccctg	gtgcgcttca	gccctgactg	cagagccttc	480
atcgtctggc	tgcccaacgg	ggacaccctc	cgtgtcttca	agatgaccaa	gcgggaggat	540
gggggctaca	ccttcacagc	caccccagag	gacttcccta	aaaagcacia	ggcgctgttc	600
atcgacattg	gcattgtctaa	cacagggaa	tttatcatga	ctgcctccag	tgacaccact	660
gtcctcatct	ggagcctgaa	gggtcaagt	ctgtctacca	tcaacaccaa	ccagatgaac	720
aacacacacg	ctgctgtatc	tccctgtggc	agattttag	cctcgtgtgg	cttcaccca	780
gatgtgaagg	tttgggaagt	ctgctttgga	aagaaggggg	agttccagga	gggtgtgcga	840
gccttcgaac	taaagggcca	ctccgcggct	gtgcactcgt	ttgctttctc	caacgactca	900
cggaggatgg	cttctgtctc	caaggatgg	acatggaaac	tgtgggacac	aratgtggaa	960
tacaagaaga	agcaggaccc	ctacttgcg	aagacaggcc	gctttgaaga	ggcgccgggt	1020
gccgmgccgt	gccgcctggc	cctctccccc	aacgccagg	tcttggcctt	ggccagtggc	1080
agtagtattc	atctctacaa	taccggcgcg	ggcgagaagg	aggagtgtct	tgagcgggtc	1140
catggcgagt	gtatcgccaa	cttgtcctt	gacatcactg	gccgctttct	ggcctcctgt	1200
ggggaccggg	cgtgtcggct	gtttcacaa	actcctggcc	accgagccat	gggtggaggag	1260
atgcagggcc	acctgaagcg	ggcctccaac	gagagcacc	gccagaggct	gcagcagcag	1320
ctgaccagg	ccaagagac	cctgaagagc	ctgggtgccc	tgaagaagtg	actctgggag	1380
ggcccggcg	agaggattga	ggaggaggga	tctggcctcc	tcatggcact	gctgccatct	1440
ttcctcccag	gtggaagcct	ttcagaagga	gtctcctggt	tttyttactg	gtggccctgc	1500
ttcttcccat	tgaactact	cttgtctact	taggtctctc	tcttcttgct	ggctgtgact	1560
cctccctgac	tagtggccaa	ggtgcttttc	ttcctcccag	gccagtgagg	tggaatctgt	1620
ccccacctgg	cactgaggag	aattggtag	aggagaggag	agagagagag	aatgtgattt	1680
ttggccttgt	ggcagcacat	cctcacaccc	aaagaagttt	gtaaatgttc	cagaacaacc	1740
tagagaacac	ctgagtacta	agcagcagtt	ttgcaaggat	gggagactgg	gatagcttcc	1800
catcacagaa	ctgtgttcca	tcaaaaagac	actaagggat	ttccttctgg	gcctcagttc	1860
tatttgaag	atggagaata	atcctctctg	tgaactcctt	gcaaagatga	tatgaggcta	1920
agagaatatc	aagtccccag	gtctggaaga	aaagtagaaa	agagtagtac	tattgtccaa	1980
tgtcatgaaa	gtggtaaaag	tgggaaccag	tgtgctttga	aaccaaatta	gaaacacatt	2040
ccttgggaag	gcaaagtttt	ctgggacttg	atcatacatt	ttatatgtgt	gggacttctc	2100
tcttcgggag	atgatatctt	gtttaaggag	acctcttttc	agttcatcaa	gttcatcaga	2160
tatttgagtg	cccactctgt	gcccaataaa	atatgagctg	gggattaaaa	aaaaaaaaaa	2220
aaaaaaaaaa	ctcga					2235

<210> 37

<211> 2971

<212> DNA

<213> Homo sapiens

<400> 37

gacgtgagga	gcgttccatt	tggccagtgg	tgggcgggtg	ccacagctgg	tttagggccc	60
cgaccactgg	ggccctctgt	caggaggaga	cagcctccc	gcccgaggag	gacaagtgc	120
tgccaccttt	ggctgccgac	gtgattccct	gggacgggtc	gtttcctgcc	gtcagctgcc	180
ggccgagttg	ggtctccgtg	gttcaggccg	gctccccctt	cctggctctc	cttctccgc	240
tgggcgggtt	tatcgggagg	agattgtctt	ccagggtcag	caattggact	tttgatgatg	300
tttgaccag	cggcaggaat	agcaggcaac	gtgatttcaa	agctgggctc	agcctctggt	360
tcttctctcg	tgtaatcgca	aaacccattt	tggagcagga	attccaatca	tgtctgtgat	420
ggtggtgaga	aagaaggtga	cacggaaatg	ggagaaatc	ccaggcagga	acaccttttg	480
ctgtgatggc	cgcgtcatga	tggcccgca	aaagggcatt	ttctacctga	cccttttctt	540

20

catcctgggg	acatgtacac	tcttcttcgc	ctttgagtgc	cgctacctgg	ctgttcagct	600
gtctcctgcc	atccctgtat	ttgctgccat	gctcttcctt	ttctccatgg	ctacactgtt	660
gaggaccagc	ttcagtgacc	ctggagtgat	tcctcgggcy	ctaccagatg	aagcagcttt	720
catagaaatg	gagatagaag	ctaccaatgg	tgcggtgccc	cagggccagc	gaccaccgcc	780
tcgtatcaag	aattttccaga	taaacaca	gattgtgaaa	ctgaaatact	gttacacatg	840
caagatcttc	cggcctcccc	gggcctccca	ttgcagcatc	tgtgacaact	gtgtggagcg	900
cttcgaccat	cactgcccct	gggtggggaa	ttgtgttgga	aagaggaaact	accgctaact	960
ctacctcttc	atcctttctc	tctccctcct	cacaactctat	gtcttcgcct	tcaacatcgt	1020
ctatgtggcc	ctcaaatctt	tgaataatgg	cttcttgagg	acattgaaag	aaactcctgg	1080
aactgttcta	gaagtcctca	tttgcttctt	tacactctgg	tccgtcgtgg	gactgactgg	1140
atttcatact	ttcctcgtgg	ctctcaacca	gacaaccaat	gaagacatca	aaggatcatg	1200
gacaggggaag	aatcgcgctc	agaatcccta	cagccatggc	aatattgtga	agaactgctg	1260
tgaagtgcgtg	tgtggccccct	tccccccag	tgtgctggat	cgaaggggta	ttttgccact	1320
ggaggaaagt	ggaagtcgac	ctcccagtac	tcaagagacc	agtagcagcc	tcttgccaca	1380
gagccagacc	cccacagaac	acctgaactc	aaatgagatg	cgggaggaca	gcagcactcc	1440
cgaagagatg	ccacctccag	agccccaga	gccaccacag	gaggcagctg	aagctgagaa	1500
tagccctatc	tatggaagag	acttttggtt	gtgtttaatt	agggctatga	gagatttcag	1560
gtgagaagtt	aaacctgaga	cagagagcaa	gtaagctgtc	ccttttaact	gtttttcttt	1620
ggcttttagt	caccaggttg	cacactggca	ttttcttgct	gcaagctttt	ttaaaattct	1680
gaactcaagg	cagtggcaga	agatgtcagt	cacctctgat	aactggaaaa	atgggtctct	1740
tgggcccctgg	cactggttct	ccatggcctc	agccacaggg	tccccttggg	ccccctctct	1800
tcctccaga	tcccagccct	cctgcttggg	gtcactggtc	tcattctggg	gctaaaagt	1860
tttgagactg	gctcaaatcc	tcccaagctg	ctgcacgtgc	tgagtccaga	ggcagtcaca	1920
gagacctctg	gccaggggat	cctaactggg	ttcttggggt	cttcaggact	gaagaggagg	1980
gagagtgggg	tcagaagatt	ctcctggcca	ccaagtgcc	gcattgccc	caaatecttt	2040
taggaatggg	acaggtacct	tccacttggt	gtatttatta	gtgtagcttc	tcctttgtct	2100
cccatccact	ctgacacct	agccccactc	ttttccatt	agatatatgt	aagtagttgt	2160
agtagagata	ataattgaca	tttctcgtag	actaccaga	aactttttta	atactgtgct	2220
cattctcaat	aagaatttat	gagatgccag	cggcatagcc	cttcacactc	tctgtctcat	2280
ctctcctcct	ttctcattag	ccccctttta	ttgtttttc	cttttgactc	ctgtcccat	2340
taggagcagg	aatggcagta	ataaaagtct	gcactttggt	catttctttt	cctcagagga	2400
agcctgagtg	ctcacttaaa	cactatcccc	tcagactccc	tgtgtgaggc	ctgcagaggc	2460
cctgaatgca	caaatgggaa	accaagggcac	agagaggctc	tcctctcctc	tcctctcccc	2520
cgatgtaccc	tcaaaaaaaa	aaaaaatgct	aaccagttct	tccattaagc	ctcggtctgag	2580
tgagggaaag	cccagcactg	ctgccctctc	gggtaactca	ccctaaggcc	tcggcccacc	2640
tctggctatg	gtaaccacac	tgggggcttc	ctccaagccc	cgctcttcca	gcacttccac	2700
cggcagagtc	ccagagccac	ttcacctctg	gggtgggctg	tggccccag	tcagctctgc	2760
tcaggacctg	ctctatttca	gggaagaaga	tttatgtatt	atatgtggct	atatttccta	2820
gagcacctgt	gttttcctct	ttctaagcca	gggtcctgtc	tggatgactt	atgcygtggg	2880
ggagtgtaaa	ccggaacttt	tcatctat	gaaggcgatt	aaactgtgtc	taatgcaaaa	2940
aaaaaaaa	aaaaaaaa	aaaaaaaa	a			2971

<210> 38

<211> 1163

<212> DNA

<213> Homo sapiens

<400> 38

ccacgcgtcc	gccaaggggtg	ctgattaggg	aatggttatg	gactaggagt	atcagtaaca	60
atgggttagaa	agtggctaac	atttgttgag	cacctgctgt	gtgcctggcc	ccggctggga	120
gccttcgtgc	ccagagtgc	ccgctctgaa	tgcagttctt	tgcctcatc	aaactgggga	180
gtgggaggca	gagctgcaca	actcacaggt	gccgagctca	agactcactc	ctgggtctgc	240
ctgggctggg	ctgtgcttgt	tgcctctgtg	gccaacacac	gcgcaccttt	cacctgaaag	300
ccaggatccg	cagaacgttc	cccaggagg	tcattgtttg	gcactatgat	ttgtctcttc	360
ctaaaaaggt	gatagagtta	cactggagag	agcagcatcc	aggtgcagca	gggatgggcc	420
tggggctcac	gggcaggggc	tctgtgtccg	gctggggcct	gggttctctc	gctgcacctg	480
tgtgtcagaa	gcactcagta	aatctttgct	gatgaaggat	gacaggatat	aggacatgat	540

21

```
gcttgctgct gcattgcctg caatcctgga tgaatgccca ggttggtttt gctccccgtc 600
gggtggatgt gacgttagct gtgatgttag gtccctggct ttaaaatagc acggaactgg 660
gaattgaggg agcagtttgg gcagaaagga cagccccgca gaggcctgga gctgagcagt 720
gcgggcgacc caggagcagt gagtgttcc gtcacagcct tcacgcacc ctgtggtcct 780
cataaagggg atggaatcta cgaatttagt tttcccagcc tccttaaaaa ctcatctcatg 840
ccaggggagcagg tggctcacac ctgaaatccc accactttgg gaggtgagg caggctgatt 900
acttgaggtc agggagtttga gaccagccta gccaacatgg tgaaaccccg tgtctactca 960
aagtacaaaa aaaaaaatta gtcagacgtg gtgtcacgca cctgtaatcc cagctctttg 1020
ggaggctgag gcaggagaat cacttgaacc caggaggcag aggttatagt gagccagtat 1080
tgccgcaactg acctccatct gggcaataga gtgagaccct gtctcaaaaa aaaaaaaaaa 1140
aaaaaaaaaa aaaaaaaaaa aaa 1163
```

<210> 39

<211> 1932

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1624)

<223> n equals a,t,g, or c

<400> 39

```
ggcagagacc agggccctgg gccgggcgct gagggcgggc cctctgggca gggcccgggc 60
ggggctgggt gggccgcccc tgctgctgcc gtccatgctg atgtttgcgg tgatcgtggc 120
ctccagcggg ctgctgctca tgatcgagcg gggcatcctg gccgaratga agccccctgcc 180
cctgcacccg cccggccgag arggcacagc ctggcgcggg aaagccccca agcctggggg 240
cctgtccctc agggctgggg acgcggaact gcaagtgcgg caggacgtcc ggaacaggac 300
cctgccccgg gtgtgcggag agccaggcat gccccgggac ccctgggact tgccgggtgg 360
gcagcgggcg accctgctgc gccamatcct cgtaagtgcg cgttaccgct tcctctactg 420
ctacgtcccc aaggtggcct gctctaactg gaagcgggtg atgaaggtgc tggcaggcgt 480
cctggacagc gtggacgtcc gcctcaagat ggaccaccgc agtgacctgg tgttcttggc 540
cgacctgcgg cctgaggaga ttcgctaccg cctgcagcac tactttaagt tcctgtttgt 600
gcggggagccc ttggaacgcc tcctctctgc ctaccgcaac aagtttggcg agatccgaga 660
gtaccagcaa cgctatgggg ctgagatagt gaggcggtag agggctggag cggggcccag 720
ccctgcaggc gacgatgtca cattccccga gttcctgaga tacctgggtg atgaggaccc 780
tgagcgcatg aatgagcatt ggatgccgt gtaccacctg tgccagcctt gtgccgtgca 840
ctatgacttt gtgggctcct atgagaggct ggaggctgat gcaaatcagg tgctggagtg 900
ggtaggggca ccacctcacg tccgatttcc agctcgccag gcctggtagc ggccagccag 960
ccccgaaagc ctgcattacc acttgtgcag tgccccccgg gccctgctgc aggatgtgct 1020
gcctaagtat atcctggayt tytccctctt tgccctacca ctgcctaagc tcaccaagga 1080
ggcgtgtcag cagtgacct ggggtgtggg ccagcagctg gtggggactg gtttcaacgc 1140
cagctttctg tgctttctgc tgtcattcgg agaaactctg gctctggggc ttggggcttc 1200
tcaggatcct ggatggcaga gactgccctc agaarttcct tgtccagggt gggcaccac 1260
agtgactcag aggacagggc taggcaggag acctgctgct cctcattggg gggatctctt 1320
ggggggcaga caccagtttg ccaatgaagc aacacatctg atctaagac tggctccaga 1380
ccccgggctg ccaggattat gcagtcact tggctctact taatttaacc tgtggccaaa 1440
ctcagagatg gtaccagcca ggggcaagca tgaccagagc cagggaccct gtggctctga 1500
tccccattt atccacccca tgtgcctcag gactagagt agcaatcata ccttataaat 1560
gacttttctg cttttctgct ccagtctcaa aatttcctac acctgccagt tctttacatt 1620
tttnccaagg aaaggaaaac ggaagcaggg ttcttgctct gtagctccag gacccagctc 1680
tgcaggcacc caaagaccct ctgtgccag cctcttctct gagttctcgg aacctctctc 1740
ctaattctcc ctctctctcc cacaaggcmt ttgaggttgt gactgtggct ggtatatctg 1800
gctgccattt ttctgatgca tttatttaaa atttgtactt tttgatagaa ccttctgaag 1860
ggctttgttt tcctaatagc tgacttttta ataaagcagt tttatataaa aaaaaaaaaa 1920
aaaaaaaaaa aa 1932
```

<210> 40
<211> 881
<212> DNA
<213> Homo sapiens

<400> 40
gaattcggca cgagggaacc cagaagatgc tgectctect gatcatctgt ctctgcctg 60
ccattgaagg gaagaactgc ctccgctgct ggccagaact gtctgccttg atagactatg 120
acctgcagat cctctgggtg accccagggc caccacaga actttctcaa agtattcact 180
ccttgttcct agaggataat aattttctca aaccttggtta ccttgatcgt gaccatttgg 240
aagaagaaac agccaaattc ttactcaag tacaccaagc cattaaaacg ttacgagatg 300
ataaaacagt acttctggaa gagatctaca cgcacaagaa tctctttact gagaggctga 360
ataagatata tgatgggctg aaggagaagg gagccccacc cytctccatg aatgccttcc 420
cggctccatc tcctacttgc accccagaac cccttggtc tgctgcctc cccagcacct 480
cagtttctct accttctcac ctccctggca gcctgcaatg agtctgtgc caggaaccgg 540
cggacctccc tgtgggctgt gactctcagc agtgctctac tcctggccat agctggagat 600
gtttctttta ctggcaaagg aagaaggagg cagtaaagga acagggcagc ccgcatgtct 660
tccagaagtg aacagaggcc gcagctacca ccgtcacaaa gttcaactcat ctctgggtcc 720
cgggtgacccc atccccccat accctccatc ctgggtcctg gggcccaaaa gctctgaggc 780
ctaggagact gcgctgtctc gtgggttgc tactctaca cctttgtaaa gagtctcttc 840
attaaaaccc ctcttcataa aaaaaaaaaa aaaaaactcg a 881

<210> 41
<211> 1932
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (2)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1022)
<223> n equals a,t,g, or c

<400> 41
cncggcgagg ctcggctcat gccccggggc gcggggcaca caggccggcc ggcagccgct 60
gggaaatagg cccccggggg cggtggcggc ggccggggcca tggcgcgagg accccggggc 120
cgggcccgtt cgggggagga gttctcttc gtcagccgcg tggtgaaata cctgctcttc 180
ttcttcaaca tgcctctctg ggtgatttcc atggtgatgg tggctgtggg tctctacgct 240
cggctaataa agcatgcaga agcagcccta gcctgcctgg cagtggaccc tgccatcctg 300
ctgatcgtgg tgggtgtcct catgttctct ctcaccttct gtggctgcat tgggtccctc 360
cgcgagaaca tctgcctcct gcagacgttc tccctctgcc tcaccgctgt gttcctgctg 420
cagctggcgg ctgggacccg gggcttcgtc ttctcagaca aggctcgagg gaaagtgagt 480
gagatcatca acaatgccat tgtgcactac cgagatgact tggatctgca gaacctcatt 540
gattttggcc agaaaaagtt tagctgtgtt ggagggattt cctacaagga ctggtctcag 600
aacatgtatt tcaactgctc agaagacaac ccagtcgag agcgtgctc tgtgccttac 660
tcctgttgct tgcctactcc tgaccaggca gtgatcaaca ctatgtgtgg ccaaggatat 720
caggcctttg actacttggg agctagcaaa gtcacttaca ccaatggctg tattgacaag 780
ttggtcaact ggatacacag caacctattc ttacttgggt gtgtggctct aggcctggcc 840
atccccccagc tgggtgggaat tctgctgtcc cagatcctag tgaatcagat caaagatcag 900
atcaagctac agctctacaa ccagcagcac cgggctgacc catggtactg agaatccatc 960
ctgcacctcc tcaccatgga aactggcaag cctcataaac gaacagcagt ggggtgtgaa 1020
ancagacca aatggagatt tggattccag cccccagtg acagcccagt gggagaagc 1080

23

aaactccaga	tgggcagaag	gcaggggtgca	caggtggctc	cagtctcagg	aggatgcgcc	1140
tcctctcccc	catcccagcc	ctcagcattg	tgccagagtg	atacccttaa	gtgttttgggt	1200
ttatgttttc	agttttgttt	gggaacacag	agttgcacag	agagttgggg	gtactgtctgc	1260
tgccttttca	ccgaggcact	gccaccacca	gctctascag	ggatgctcct	gagcttggcg	1320
gacatactta	gacctaacg	tgccagtgag	acctggctgt	ggagagtagc	actggcagcc	1380
ctgcctggac	tccacttggc	atgataccag	ctccagaagg	gaagggagtg	gagcaggcag	1440
tgaggagaga	gcctgggggt	cggctgggga	cagccgtatg	tgctaggtag	gagtgagggg	1500
agatatgttt	accaaagtgc	tgctctgcca	tcctcccagg	tagtcagagt	gagctacatc	1560
ctgccccgcc	ttcatttcca	tggaaacatg	gcagctagga	cacggggtag	aacagcagcc	1620
aaattcttcc	ccacctccct	tacttcgaaa	aaaagtttgg	aaccctggtc	cctatactct	1680
gcagtcagaa	gtgggactga	gccatacatg	cccttgaatt	cctccctgtc	tggccctccc	1740
tctccagcaa	gcaggggttt	ctttaacttg	gcagtggtga	gaggagaagt	ggtaaacacc	1800
ccacccatt	cccctgcac	ggagctcagt	attcctacag	ggtaagaggt	aggaatcttg	1860
ctgggacgag	gggagccaga	agtggcaata	aaagcgtgtt	gacctggaaa	aaaaaaaaaa	1920
aaggcgggcc	gc					1932

<210> 42
<211> 1164
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (582)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (592)
<223> n equals a,t,g, or c

<400> 42						
ggcacgagct	tgtgtgtcac	cagcctcctg	atctgccagg	gtctgctctg	ggttggcact	60
gaccaggggtg	tcacgtcct	gctgcccgtg	cctcggtctg	aaggcatccc	caagatcaca	120
gggaaggga	tgtgtctact	caatgggcac	tgtgggcctg	tggccttcc	ggctgtggct	180
accagcatcc	tggccccga	catcctgcgg	agtgaccagg	aggaggtga	ggggccccgg	240
gctgaggagg	acaagccaga	cgggcaggca	cacgagccca	tggccgacag	ccacgtgggc	300
cgagagctga	cccgaagaa	gggcatcctc	ttgcagtacc	gcctgcgtc	caccgcacac	360
ctccccggcc	cgctgtctc	catgcgggag	ccggcgctg	ctgatggcgc	agctttggag	420
cacagcgagg	aggacggctc	catttacgag	atggccgacg	accccgacgt	ctgggtgcgc	480
agccggccct	gcgcccgcga	cgcccaccgc	aaggagattt	gctctgtggc	catcatctcg	540
gcgggcagg	ctaccgcaac	tttggcagcg	ctctgggcac	antgggaagc	angccccgtg	600
tggggagacg	gacagcacc	tcctcatctg	gcagtgcct	tgatgtata	gcgcctcccc	660
tctcccctca	gagggcacag	ctgcaggcct	gaccaaggcc	acgcccggct	ctcgtgtctt	720
aggacctgca	cgggacttgt	ggatgggcct	ggactctcca	gaaactactt	ggccagagc	780
aaaggaaaac	ctcttgtttt	aaaaaaattt	ttttcagagt	gttttgggga	ggagttttag	840
ggcttgggga	gagggaggac	acatctggag	gaaatggcct	tctttttaaa	agcaaaaaac	900
acaaaacctc	acaactgcct	ggcaagccca	gtatcacttg	tttgggcct	agcgggactc	960
caaggcagcc	acacgcccct	cctggaagg	tgtgtgcgtg	tgagtgtgtg	cgagtgtgtg	1020
ggctgtgtg	tgaatatcta	taaataagta	tatatgtgt	atattatatg	tgtataaata	1080
aagtctgtac	atattggagc	tctgggagat	gctggaataa	aagacaagag	ttacatctgg	1140
acttgaaaa	aaaaaaaaaa	aaaa				1164

<210> 43
<211> 1105
<212> DNA

<213> Homo sapiens

<400> 43

gaattcggca	cgagaacaaa	ttgaaacat	ctggcatga	acttttattt	gttaagaggt	60
tttctaata	tgattcaatc	tctttgctta	ttctaataat	gttcagattt	tccacttctt	120
gagtcgaatt	ggttaatttat	ttgtttctag	gaatttgccc	atttcaccta	gtttaccta	180
tttttgacat	ataaaattat	atatggaaat	ttctaaaata	tttaaaaatt	tctgtaattg	240
caatagtaat	gtccctctt	ttgttacca	ttgttattt	gaatcttctc	cttttttttg	300
tcaatctagc	taaaaatttg	tcaattttgt	tcgtctcttc	aaaaaaatat	acgtttgtct	360
tcatgatttc	tctawtggtt	ttccatcyat	atttcatttg	aatacatttt	taaacyttay	420
ctttattatt	tcattctctc	tgggagcttt	gggtctcatt	tttttttcc	gataatctag	480
ttgtttattg	tataagatta	agtattttat	tgaaatctgt	atgttcttta	atgtaggcat	540
tcaactactat	aaatttactt	ctcaggagca	tctctgccgc	attccatggt	ttagtatggt	600
gtgttttaaat	ttgtattcat	aactagaggg	aaacagaggt	gacggagaaa	aagacgtaca	660
aatatcatcc	acttgcaaa	tatagatttg	tttgtattgk	ratatgaatr	aaaaatttac	720
gagacagata	agaaaatttg	aacactgacc	attgatgcag	ttacagttaa	ttttaaaatc	780
aagggttaata	acatttttagt	tatttttaa	aatgatagta	atttagagat	gtattctgaa	840
tggttttaaa	tgaaaagata	tgccctggat	ttcttccaaa	atgaatcttg	taggttggga	900
agaaaatgag	aacatagtag	aaacaagact	gacaatgagt	tggttagggt	gggcaatgag	960
tacactaaag	cttattttat	cttattttac	tgtatatact	gttaaagctt	gcattatttt	1020
cataaatgca	tttgctaagt	gcaactgtta	tcaataaaag	tggattgggc	tctaaaaaaa	1080
aaaaaaaaaa	aaactcgagg	ggggg				1105

<210> 44

<211> 1262

<212> DNA

<213> Homo sapiens

<400> 44

cagcatgtac	ccagttgttc	tttctctga	gaaagcaaaa	tgccatgat	ttcttataat	60
ccaggctgcc	acgtttacct	tgtaaaatca	atacttaatt	tttagatttt	tatattatct	120
tttctctgta	agcaagactt	ctaaattatg	gctataatat	cttttgaatt	gttggtctta	180
atgaatcttc	caactgtaaa	ctcatctaat	ttcaaaactta	tcatacctga	ggatgtaaca	240
ttgtctcttg	tttctcatct	tgatattacc	gtcaatcatt	ttgtatttct	gagtacattt	300
gaacttgctg	gagtaataga	gggaaaacct	ctgcttgatt	ctaaatcaga	tctttgtcct	360
atactcggac	aattatggtt	tcatatttta	ttatttttta	ttttctgggt	ttaacaaatg	420
agataacatt	ttagacataa	tatttgtaaa	catcttgact	tatttcagca	ttttcttttt	480
ttgtgtatct	tcagagagtt	tggtgaaagt	agcaatttcc	aagtaatttt	aaattattga	540
agtctactag	cacgaaaagg	caaattctta	ggatatttta	aaaatgttgt	ttaataatca	600
aactcatctt	aaaaaatggt	catcagactc	tgtctttgat	gcacattttg	ccaaaagaga	660
gccttatttc	tgtgaaagaa	atacagtagt	tactttggga	tttactaaag	taaaactggt	720
actttaaggc	acagagcaga	tatagaatcc	ccctctctcc	ccactcctag	tgactgggtat	780
tctacattaa	tatttatctt	ccatgcatag	tgtacttgag	ggaaaaaac	aataactctt	840
aattgtttta	tatcaacaa	taaaatcctg	tgtatcagtg	actgtcaata	gatggctttc	900
tgtttaaaaa	ctgaagctac	tccagaagta	ggaattaatt	tatttagtaa	acaaagtcag	960
tcaaacacaga	gccatgtcct	ggggaactgt	caaaagaatg	gttcctaagg	gccagagggc	1020
acatccactg	gtagatgaca	gaacaacat	acttcagatg	gcaaaaccgg	tcagtttggt	1080
ttgcgttggtg	tgccatctct	ctttctgtgt	gcttcagctg	aattaagtgc	ttggagagct	1140
caaatagttc	aagatagcca	agatgaccaa	ttctgccagg	tggcaagcct	gatcttgcaa	1200
tttgatttaa	aataaagaac	attccccaag	aacagtttgt	tgcaaaaaaa	aaaaaaaaaa	1260
aa						1262

<210> 45

<211> 517

<212> DNA

<213> Homo sapiens

<400> 45
gaattcggca cgagtgcact tccaccagct atgtatgaga cttcccatg ctccacatct 60
ccagtatttt atgtggtcag tccttttgtt ttgtgtcatt ttgtggata tgaaatggca 120
tctcagtgtg gcttttcatt atatttcctt gatgactaat ggtattcctt caccctttca 180
gtgcttattg gccattcatg tatctttgtt ttttgtgtag cacttcaggt cttttgccc 240
tagatttagt ggggttgatt ctctttatta atgatttga gggatgttat atatattctg 300
gacacaagat tattgttaga gatacgtaact tcagatattt tctcccagtc ttagcttgc 360
ctaattatta ttattattat tatttgagat gaagtctcac tctgtcgccc aggcagaggt 420
tgcaaggggc cgagatagca ccactacact caagcctggc tgacagagtg agactctgtc 480
tcaaaaaaaa aaaaaaaaaa aaaaaaaaaa aactcga 517

<210> 46
<211> 858
<212> DNA
<213> Homo sapiens

<400> 46
agaaaaaatc ctacatggat attggttaga aagagagaaa ggaagtggcc agtgtcccgt 60
ggcctcttcc accttctgga ttgttgaagc tggggcctgg aggggatggc cctgccactc 120
agcagggggc actaatggga ccaagctaac ctgtccagtg agaattcctgc agggagacct 180
gaggggtacca ggaaagtga ggggaaggcc cgggaaatgg agagagctgg tctggagggg 240
aggagcaagc cgcgtggggc aggccatgtg ccttttgcct gggggagtat tactcatttg 300
gagctgtgctg tctggaacgc ctgcctcaca cacaaggac tggggcagat gtaagttctc 360
tgcaagcaagc aagcgcacag ctgagagtaa cttagaaagc acccagctaa tgctggcatc 420
ccagatcgac cccctcctcg ctgaatgttg gcattctctg gcctcagttt cctcatctgt 480
aatgggggtg gataagaaat gtgtacacac ctcccgggca gtggggagga ttaaaactgtg 540
ctctgacacg atccggggcat gtccaggggtg gtatctgcag taaaccgcgc tcggaaaatg 600
gcgggcgcatc agggccagcg gtgggagctc tccgtgcttg gcttgacgcc attgtggagg 660
tggaggaggg gctgcaagac tctgagcagg aagacccgc aaagcaggaa agcagagcca 720
gagttggggg ccagccgcag aaacgagagc ccccgctgact ttgaggcacc ctttgagag 780
ggcaggaagc aggaagggtg aattttctcc aaaacccaag aggcagagtg accccacatg 840
ataactgagt ttctcgag 858

<210> 47
<211> 6107
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (5749)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (5892)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (5896)
<223> n equals a,t,g, or c

<220>
<221> SITE

26

<222> (5906)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (5957)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (5966)
<223> n equals a,t,g, or c

<400> 47
gcagttagtt ccttgatgtc agtagtgggc taaaggcagc ttactgtgtg tttgctggag 60
ctttcactca gccaaagtgt agagtcagga aaccattga ggcaatggcg tcaaatgggtg 120
tttcacaaga atgagccatt cagtccttgc tcactatata tttaatatgt tattattgtt 180
gttattgtta ttattaattg gctttctgta ttctatgcct tttatttata aagacactaa 240
gaaaacccat gtttgtaatt ttaataacat ttttcccatc ttgtaatatc cagagctact 300
ttataaattc tctgaaccaa aagtatttcc ctcagtgtat ctcttctccc ccagccccta 360
ttgggaaaaa ttaccagta tagttcaggt tatgaggagg atcagccaca caatccagtg 420
cttcagtttg aaaatgtaaa attctaacc taaagtaggg ttgggtgaaa tttcagacaa 480
agcaaaccga gcaggtataa aaagtagtat aaatacaaat ctgtaagtta tttttgaatt 540
ttctgaactt ttttctaaga gattacatag gagactaaag aaatctatct gttcaagttc 600
taattaggat gattgttaat actgcactgt ggatgaagtg gcgactggct tgtgtgctga 660
cttctgtgtt ttagcaagag gtttattgtt atcaaatgct aattggcaat gccaaagtcac 720
tgggaccaat tttctgtttt ataatatcta agtttagaac agaatatata cctgaactgt 780
agtggtttga tcggatggag acagaaaaacc cgatttttat tctcataaat tttgtggtta 840
tttatacaag ggctgtgcta tgctaccata ttcttgttca ataataatag gtttgttgtt 900
ttttttacat tgttaaattg tccttaccct taaagggtcaa tgttaagtac aacattctga 960
aaatacaatt tggctacgaa gagtattcat ctcttttgaa gctcagtggt tgatatttgt 1020
gctaataatg caatttctcg attcctgtta caagttatag ctacatatgg gagagactca 1080
gtgagccagc aaaggccata gaaacaacaa tttattaaat gtattttatg cagaaggacc 1140
taaataaact gtgagccacc ttttcttctt tatattgtta catttaagt tttctgtctt 1200
cagcaactca cattaatgct tggagcttat ctcttctctc ctctctctct ctctctctct 1260
ctgtgtgtgt gtgtgtatgt gtgtgtgtgt gtgtgtgtgt ttccttattg tcattccatt 1320
atatatccac accaactatg gtgacgataa ttcaaagtca tattttgcct ctaagcttga 1380
tcattgtacc tttatgatta aagtatcatg ttatttagcc aatgcaaate tgttttaaaa 1440
caaatagttt aaaaaagaa caagttttta agggctttat tatagaagaa gtaataatga 1500
aggactttcc ttctctctcc cctttctccc cctccctgcc tccctctctc ccttccatct 1560
ccccctctc cctgccttct ttgtttctcc ttccttatt cctccctccc tcccttctcc 1620
cttctctctt ttcttccatt catccttctc tgccttttat ttttattttt tgtaatatca 1680
catgtgtgtg agtttggaa tttattctag tgcatttctt gctcatcaga acctcagcta 1740
atctacctag gaaaaatag atcaaaggaa atgagaaagt tgtatctgag tccctccaga 1800
actaagataa ttctttttga ccatttaagc ctttataaat gcgttttgac catttaagcc 1860
tttataaatg cttgttttag gaaagtgaat ctggttagat catcaacaaa taatgaccag 1920
gacaaaacga ttttaataat aaagtctcaa atcaccatgg ttatacat tccaccagaaa 1980
tagtaatctt acaatttttc atttttctga tgaagatttc tgttccaata tctgtttcct 2040
aatagatttt ttaaatatag tagctttcct ctgctttatg accacaggtt ttatccctaa 2100
ccgagacagc tgtcttatat ctgcatgcct tagactgtgt ggagggactc catgaagaaa 2160
gaccataggt tagaaaaata actcatagta tataccctag taagtgggtt agtagaatct 2220
cataacatgt attaaaaaga ggttttcttc tctgcttgtt tgtgtcacta gagcaaaatt 2280
gtagagataa tgctcataat gcagtaataa tcagaataat ctacaatatc atttgggat 2340
gggtcccagggt ccagtgctc tagttacttt acttcttttt ttttttttga gatggagtct 2400
tgctctgtct ctcaggctag agcagtgtgc gatctcagct cactgcagcc tccacctccc 2460
aggttcaagc gattctcctg cctcagcctc ccaagtagcc aggtattacag gcacctccca 2520
ctaggcccg ctaatttttt ttgtattttt ttagtagaga tggggttttg ccatgttgcc 2580
caggctgggt tcgaactcct aacctccagt gatccacctg cctcggcgct ccaaagtgtc 2640

aggattacag	gcatgagcca	ccacatccgg	cctaattact	tctttaatcc	ccatttat	2700
ttatgccatt	ctagccctcat	ttattaataa	aattatgttt	ttactttctc	tttcaggaaa	2760
ttttttaaat	taatatattta	tatctagatc	taatgctatg	gaaaagtgcc	tttttatcat	2820
ttataatttc	atttttcact	atttccaaaa	acacataaac	aaatagtttc	agtaggtccc	2880
agcttttact	ttttccattt	aaaccttctt	ttctccattt	cttccctttg	gcttaagaat	2940
aaaagaaaag	gtacattgct	agaattgttt	ctttgggaga	gggtaaaaga	ttacagaatt	3000
agactgttca	gcctttatat	aaactaaatt	tgtcttcac	tcaaccagct	aatggtaggt	3060
cttatctgaa	tactcatgag	aatttttagca	tctgtgaaac	tccatgcacc	agatgtgtgt	3120
aaatttcagg	aagaaagtgt	tgaagcattt	ttctctgatg	ttaattagat	ggaaaataat	3180
cactaaaaca	tagtttaggt	aaagcctgat	tatgccactt	ttttttaact	agacagggca	3240
aagttgttta	tgtagtgta	cttcttgtct	atcctcagtt	aatttaccta	gacaaaaagt	3300
gtcaaaaggaa	atgagaaaaa	gggtatatct	gactccctcc	agacctaaaga	taattccctt	3360
tgatcagata	cagtcagatg	gagtgccctg	gtttttgtta	attttgctc	tattccagct	3420
ccttaccaca	gcggtgggtc	ttaaagaaag	gatcatcagc	aacaggtcag	gatagtctta	3480
cctttgggat	agggtctgtt	tcccctgtct	agtattttctg	tgactgttag	tggcactgag	3540
gactgcaaac	ttttatgcaa	tattcttaat	accttattga	tattatgcac	tttaatacatt	3600
ccaaagaagc	caagaatgct	gtatagtgat	gattccctcc	taatgaattc	atcttaacta	3660
tttagaatgt	tatgtccctt	ttcttttggga	tagccaactt	gggtataaatg	ttatatggat	3720
ttttctaaaa	tgactatata	ggacttaaga	ctttgaaatg	taatttactt	ataaggggaa	3780
ataattatgc	tttagcacat	catttttagaa	acgtcacatt	ttagaaacat	tcagcttgct	3840
aacctacatg	tttgggaatt	cattaaaacc	agttgtctat	atattttgtg	ccatgtatat	3900
aagaacatta	caatatatct	ttttctacat	atgtagtatg	tgcaaccagt	gggtctcaga	3960
gtatggttct	cagcccacca	gctagtatca	gtatcacctg	ggaactagtt	agaaatgtaa	4020
attctttggc	cccattcccag	acatactgag	tcagaaattc	tggaaatagg	ccccgcgaat	4080
ctgttttcac	aagccctcca	gggtgattctg	atgcacactt	taaagttag	gaaccactgg	4140
gctaagactc	tggtgagata	tagagttttt	cttccactca	gactgatata	gttatacatt	4200
gttcttcacg	taaatccagc	ttaacctggg	tatctataat	cttttattgg	caaaagttaa	4260
ttctcagtag	tgccctataga	gatacagtg	attttatgta	catacacaa	tagtctaatt	4320
cttgataatt	cagttaattt	agtttggcat	tttccctacca	cttactaaaa	ggtttacatt	4380
aaatgactga	tttaaatata	taggtgcaat	gttctatgtt	tattttaatt	ggtatgacat	4440
ttaagtagct	aatataattg	accggtgcta	aagtctcctg	tttatccata	aaatgggtac	4500
attatgggca	gtgtaataca	agctttcttt	tcattgccta	gtactttacc	agcagaccac	4560
agttttgccc	tggttagacc	aacctccaga	acaaaatcat	cattccctgt	atttatattt	4620
gtatctgaga	tagtaaacaa	gatggctggc	cagggtcaaca	tggcacctta	acttattttt	4680
taaataggta	aaacttcttc	aaaagtagct	tgctttgtat	aagaactaag	ctatcagtat	4740
agatatagct	atccttgagg	cttatgtttc	agacaagaat	tatttactaa	aataaataat	4800
aaacaagata	atgcattata	caatttgggc	atttctcggt	tctcaagtgt	atgcacatg	4860
gtaaatataa	actaaccaca	agataggtag	attgattcat	ttcattttaa	tctccttggt	4920
taattcagta	cctccataat	tgttctaate	ttcttcccac	tgtttacaaa	ttaccagtta	4980
attaactcgt	gaaagaaaaa	ttcacatatc	agaataaaaa	taaatgtata	ctcactttat	5040
aaaaatcacc	actgctgtct	ttccttaata	ctagcagtg	aaatgtaagt	ggcttactct	5100
acaaattttg	gtgctggcaa	atacataggc	aaactgttgg	gagctgctct	agttacattc	5160
ctcccttctt	attccctttt	tctcttcttc	actttattgc	ataacatatt	cctgtaccca	5220
aagcattcta	ccacagttct	atttgactcc	cacttgtaat	aactccttta	aaaaattcca	5280
tgtttaacca	tatgacctg	cttgcttact	catattctcc	ctccctctcc	ccttcttctc	5340
tctctcttcc	agaagtcatt	tgcttggttt	gaaatatatt	gtagggattg	cttatttatat	5400
tatttttagct	gatgaacctc	aggacaacgt	ctacacacac	acacatacat	acacgcacac	5460
aaaatctcag	ctgttggaag	gtgggcttgg	aatcagactt	ctgtgtccag	taaaaaactc	5520
ctgcactgaa	gtcatttgta	cttgagttag	tacagactga	ttccagtga	cttgatctaa	5580
tttcttttga	tctaataaat	gtgtctgctt	acctgttttc	cttttaattg	ataagctcca	5640
agtagttgct	aattttttga	caactttaaa	tgagtttcat	tcacttcttt	tacttaattgt	5700
tttaagtata	gtaccaataa	tttcattaac	ctgttctcaa	gtggtttanc	taccattctg	5760
ccatttttaa	tttttatatta	attttatttg	cttgagcaca	ctgatcaacc	actgaactgc	5820
cttcttccat	tgctctgcaa	tgatataagg	gttacatttt	tggttatatg	gctttcatag	5880
ttgggatttc	anagcnctga	taccanatat	tttcagtttg	ttctctgggg	gaatttcatt	5940
tgcatctatg	tttttancta	tctgtnataa	cttgttaaat	attaaaaaga	tattttgctt	6000
ctatttgaac	atttgtatac	tcgcaactat	atttctgtaa	acagctgcag	tcaaaaaata	6060
aacactgaaa	gttttcaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaa		6107

<210> 48
<211> 703
<212> DNA
<213> Homo sapiens

<400> 48
ccacgcgtcc gcaggacatc gttttctaca tgggtggctgt gttcctgacc ttcctcatgc 60
tcttccgtgg cagggtcacc ctggcatggg ctctgggtta cctgggcttg tatgtgttct 120
atgtgggtcac tgtgattctc tgcacctgga tctaccaacg gcaacggaga ggatctctgt 180
tctgccccat gccagttact ccagagatcc tctcagactc cgaggaggac cgggtatctt 240
ctaataccaa cagctatgac tacggtgatg agtaccggcc gctgttcttc taccaggaga 300
ccacggctca gatcctggtc cgggccctca atcccctgga ttacatgaag tggagaagga 360
aatcagcata ctggaaagcc ctcaagggtg tcaagctgcc tgtggagtgc ctgctgctcc 420
tcacagtcctc cgtcgtggac cgggacaagg atgaccagaa ctggaaacgg cccctcaact 480
gtctgcatct ggttatcagc cccctgggtg tggctcctgac cctgcagtcg gggacctatg 540
gtgtctatga gataggcggc ctctgtcccg tctgggtcgt ggtggtgatc gcaggcacag 600
ccttggtctc agtgaccttt tttgccacat ctgacagcca gccccccagg cttcactggc 660
tctttgcttt cctgggcttt ctgaccagcg ccctgtggat caa 703

<210> 49
<211> 639
<212> DNA
<213> Homo sapiens

<400> 49
ggcacgagca ttcacaggtt acaaatgctg ctgccaaactg tcttggccaa atgactctgc 60
atcacaaacc tttccttgca tggggagggg atggatttac tcagtccaac tttgatggct 120
gcatcacttc tgccctatgt gttctggaag ctttaaagaa ttatatattg tgcctatatac 180
cttattctct acatgtgtat tgggttttta ttttcacaat tttctgttat tgattatttt 240
gttttctatt ttgctaagaa aaattactgg aaaattgttc ttcacttatt atcatttttc 300
atgtggagta taaaatcaat tttgtaattt tgatagtac aacctatgct agaattggaaa 360
ttcctcacac cttgcacctt cctactttt ctgaattgct atgactactc cttgttgagg 420
gaaaagtggc acttaaaaaa taacaaacga ctctctcaaa aaaattacat taaatcacaa 480
taacagtttg tatgccaaaa acttgattat ccttatgaaa atttcaattc tgaataaaga 540
ataatcacat tatcaaaagg ccaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 600
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 639

<210> 50
<211> 867
<212> DNA
<213> Homo sapiens

<400> 50
ggcacgagca ggtactgggt gactgcctgg ctgaggaaaa gttaactaga cacttgggga 60
aaggagatcc aaggagtaga gaggcaaaat gcctttgcat gcttttcttc ctatctcttt 120
ttctttctct cttctctact ctctcccttc ctctctttct tcttttctct tctttttttt 180
tttctctttt cccccacctc tctgcctgcc tcttcccttc cctccctctc cctcccttcc 240
ccctccctcc ctccctccct tcttcccttc ctctcttctt tcttcccttc ctctccctccc 300
tcctctctcc ctcttccct gccttcttct ctctgttctg ccaacttgcc agaaggagcc 360
caagaaaaag caccagatg cttcagtcac ctctcttagaa tcttcttttt ttttatgttc 420
agaaaagatg gaaattcatt tctgctaaag agaaaagaaa aattggaaga cagggtgaag 480
gtgaacaggg ccattataag aaagaaacaa aaatctatat tctgtctaca aggaagcgag 540
agagagaaa agagagaaga aagaagttcc aggattctaa tgtaccaaag ggatctcctt 600
tttcttgttt tgttctgaaa atttcaccaa aagagcacag gagaacatct tggctaattc 660

29

attggcgatg atgtaagaaa actgagagaa atgaaagaaa tgaagaatta ctgctgcaga	720
taatatacag ccttgaggaa agaaaggctt ttaagattat agatataaag gctattgctg	780
tattctggga taaaagaaag tctgatgtca gggaaagggg aagttggaaa aactggaaaa	840
agaaaaaaga aaaaaaaaaa aaaaaaa	867

<210> 51
 <211> 1569
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (341)
 <223> n equals a,t,g, or c

<400> 51	
gtattggcca ggctggcttc aaactcctga cctcgtgac caccacacctt ggccctccaa	60
agtcagagaga ttacaggcat gagccactgc acctggcctc aagaaaaatt atatatcacg	120
tggaatagga tagtagtctc tgcactgatt ttcgttgata atggctgttc ttcttatcac	180
cattttgcta tttctttgtc tgggctatta cagggttatt acagaaattt ccagaaagac	240
ccctgcctgt cgaatgttta cttcaagctt gagctcctgg tatattatga ggaaattata	300
tgatacccca ggagaggtct tcctttccca tgccattgta naattcctaa agtaaaatta	360
atttgccttc ttgtcaaaga aggagccaat gttgttttaa aatttttagct tgagagatag	420
gtgggggaaga aattaaatag acaagtaatc mctattcaga agagaaggga gagtcattgt	480
acgaggccca agatacttgc ccaaaaaat cgcagagaaa aactagtctt tggggctcta	540
ttttttgagt ggaacatttg agttatttaa aattagaatt ttattttggt cagattagaa	600
tttctagggt atgtcatatg tgtttttaa ttgaaagctc ttaaaactcc tattgtagtt	660
taatgtcatt atccattaat ttacataaat ctgatttga tctctatttt catcgtagac	720
tgtgtagggg caatttttcc taaagggtct gtgacatagt gctacctttt ttttaaaacc	780
tgtcttgccc aggcattatt gagtgtcccc tgggtgccagc atgtgtattt cagcactgta	840
tcaacaaatc atgatcatct tctctggcca ttgtgccctt tcagattcca aacttggttac	900
ctctcagtc ttcctacaaa cttagaaagt ctaatatctt aatgtttact tatgtagcaa	960
cctcccttcc tcccacccct aaatcctctt gtaattaatt atttcccttt ggaacttttt	1020
aaatctacaa tttccttata atatggtaac caatattaat tttcttggtc tgcgccaggt	1080
ttgattttat acaaattgtt tccagtttgg gtcattgagca caaaaccagg tattttttaa	1140
aattctatata acccttcaat gaggcagtat taattttatt aactcattaa ttcaaccaat	1200
aattcttgat tgtttactgt gttgatattt ggggtatccc caatacctga cagctgtgag	1260
caaaacaaat gccctacaca catgaggtgt acagtccagt agaaaagata aacaataaagc	1320
aaattaatag ataatatgat gtccaataag gacttcaaag gaaaataaag cagagtaaag	1380
agccagagaa tgacagttag ctgtttttca catgagtcac cagaaaaggc ctctttaaag	1440
aattgacatt tgaacagaaa aacgaatcaa gggcgtcaac tgtttattgc ttttattgct	1500
taccatttga ccaagcaatt ctacacatag gattcacctt aaaaaaaaaa aaaaaaaaaa	1560
aaactcgag	1569

<210> 52
 <211> 1196
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (590)
 <223> n equals a,t,g, or c

<400> 52	
gattgggtct gtttatgtga tagattactt ttattgattt gtatgttgaa ccagccttgc	60

atcctagggg	tgaagccgac	ttgggtgtgg	tggaataagct	ttttgatgtg	ctgctggggt	120
tggtctgcca	gtgttttatt	agggattttt	gcgtcaatat	tcatacagga	tattggcctg	180
gaattttctt	tttttgttat	gtgtctgcca	ggtttttgta	tcagggtgat	gctggcctca	240
taaaataagt	tagggagggc	tccctctttt	tctttcattt	ggaagaattt	cagaaggaat	300
gggtaccgat	ccyctttgta	cctctggtag	aatttggtg	tgaatccatc	tggctckgag	360
cttttttttt	gttggttagc	tattaattac	tgccctcaatt	tcagaacttg	ttattggtct	420
attcagggat	ttgacttctt	cctgggttag	tcttgggagg	ttgtatgtgt	gcaggaattt	480
attcatttct	tctagatttt	ctcgtttatt	tgtgtagagg	tggttatagc	atyctctgat	540
ggtagtttgt	attctgtggg	atcagtggtg	atctccctt	tatcattttt	attgtgtcta	600
tttgatttct	ctctcttktc	ttctttatta	ttctygctaa	tgggtctatg	attttgttaa	660
tctyttacaa	aaacaggctt	ctagattcat	ggatgttttg	aaaggtyttt	cgtgtctcta	720
tctccttcag	ttcttccctg	atcttagcta	tttcttgtct	tctgctagct	tttgaatttg	780
tttgcttttg	cttctctagt	tcttttaacc	gtgatgtcca	gtgtgtcaat	ttcagatctt	840
tccagccttc	tgatatgggc	atttaattgt	ataaatttcc	ctcttaacac	tgcttttagct	900
gtgtcctaga	gattctggta	cggtgtctct	ttgttctcat	tgggttcaaa	taacttcatt	960
attttctgct	taattttggt	atttaccag	cagtcattca	agagcaggtt	gttcaatttc	1020
catgtagttg	tgtgggtttg	agtgaatttc	ttaatcttga	gttctaattt	gattgcactg	1080
tggtctgaga	gacgggttaca	atttccattc	ttttgcattt	gctgagaagt	gttttacttc	1140
caattgtgtc	tcgtgccgaa	ttcgatatca	agcttatcga	taccgtcgac	ctcgag	1196

<210> 53

<211> 945

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (295)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (875)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (914)

<223> n equals a,t,g, or c

<400> 53

gaatggtgaa	atattaagt	ctttctcccc	cagggttcagg	attatgacag	ctatgtccat	60
tcacctcttc	tgtacagcat	tgtcctgtgg	aagttctggc	cagtgaata	aggcaattaa	120
aagaaataaa	atatcaaag	attggaaaga	tgtaaatgtg	tcatacttca	tagaaaacat	180
gattcataga	tatacataca	cgaatgcttt	gaattcataa	gtagattcag	ccagttgctg	240
gatataaagt	caatatacaa	aaactatttt	tatagacatg	aaacacgcaa	tgagnaaaaa	300
aatttaacca	tttttagtag	catcaaaaaa	cccccatacc	taggaatatg	aatttgtagt	360
actatttggg	atatgttgat	ggatatttat	cattttccagt	ttgggattat	tataaagaaa	420
atagccctga	acattttgta	tatatgactt	ttgggtgaatg	tagcattcat	ttctgttgat	480
tacaaactca	ggggtgaaat	tggtgagtc	taaggagagct	atagatgtat	tcaacttcag	540
ctgatatggc	taataaaat	tgcgaaaaag	attgcatcaa	gttatgctcc	catcagcaat	600
atgagagttc	ctgtttttcc	acattgtcag	caacactttg	tactgttact	ccttttaatt	660
ttagccgatt	tggtggaagg	tgtggttaata	tctcattgta	gtggccaggc	gtggtgctca	720
cgctgttaat	cccagcactg	tgggaagcca	aggtgggccc	atcacgaggt	caggagatcc	780
agaccatcct	ggctaactg	atgaaaccct	gttgctgtga	gtcccaacta	cttgggaggc	840
tgaggcagga	gaatggcatg	aactcgggag	gcgngccttg	cagtgcagct	ccagcctggg	900
caacagagtg	agantctctc	aaaaaaaaaa	aaaaaaaaac	tcgag		945

<210> 54
 <211> 488
 <212> DNA
 <213> Homo sapiens

<400> 54
 ggcacgagga gagtagaggc tattcatgta atgtctataa aaaaataaca ccaaggctgg 60
 gattacaggc atgagccact gcacctggcc agtttgctta ttttgtttgg tgccctcctcc 120
 catgggagac ctcaaggagg tatgcctgcc ccacagatgc cctggaagga cagcttgctg 180
 ctccctactca gaaccacacc tgcagacaga ggaggacaga cggacactca tttgctgagc 240
 acccatgtaa catgaactaa gagctgggtg gagacaatga acgggtggagc catcggtccc 300
 gatgtggagg gagaacagct caagaccacg gaacagcctg ctctcccgt tccctggcttc 360
 cgtgcgcttt tgtccaatca ggctttttga ccaatcggcc aggcgcgcta tgtaaatttc 420
 tgacattttc aaagctgtct ttttaataaa cttttcagtg taaaaataaa aaaaaaaaaa 480
 aaaaaaaaaa 488

<210> 55
 <211> 2860
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (753)
 <223> n equals a,t,g, or c

<400> 55
 ggcacacagg gctggcaggc ccgcggtggc tgggtgtgag gcatgaacaa attgtaccgg 60
 gtatccccca cccactctg accaccagtt cctccttgga tatcactccc cctgacaggc 120
 agcccaccca ggcttgatt tgcctctgct tccccctttt gcttttcccc catgactaat 180
 gggcaccagg tcttgctgct cctgcttctc acctctgcag tggcagcagg cccctggccc 240
 cagggtcatg ccggtcagtg ggggttgatg tgccttctc caggcctgcc ctctgtccaa 300
 gcccggagtg ggcttggtg gctccctggt ggccccagtg ggggtgccagg tgggtgccgg 360
 gggtatttag ggggtggtgt atcactgtag ggacaggctt ctgccccag cctggagagc 420
 tgttttcttc aggaaggttc tggagatgga gacttggttg cgaattcacc acaactccag 480
 gctgggaggc tgggtctctg ctctcagagc cgagacacca gggaggatag ccaggctgcc 540
 ctgcctggga attctgctgg gccgtcaaat tcaacccgca ccaacgtggg caggaggcca 600
 cagtgtcctg ccaggagcag agggctgaag gtctgcagga ggaagaccct atcctggttg 660
 ggggcacctg ctgccccccc tgccccagc gtgcctgggg ggagcacacc tgggcatgga 720
 ggagtccagg gtgctggggc acacaagaga gnggggggag aggcctggac agtaggaaga 780
 tcttggcccag ggtcctggat ccgccactct gggggtgacc ttggacaaac ctctgccttg 840
 gccctcagtc tccccatcaa ggtttttcca ttcaggaggg tttgggcat cttcagccac 900
 cctaccagcc ctgaaaagga tgtgactcct gtttctggga agtgtgtggt gtgttaggtg 960
 ggcctacagc cctgggtgtg gggagggaag gatggagaga cagcacagtg acagagccca 1020
 gactgcaggc tggagttagg gtccacttc cccgtgctg tgtgtcctgg accagtgcct 1080
 ctgaaccttg gcacttgggg cagtggatat taacatcttt ccaagcccaa ttcttggggc 1140
 atcagggcct ccggtcctct gggaggtggc aggtcctcag attggagatg ccatgggggg 1200
 gggaggtgcc tctccttttg agggatgga agtgagaca ggagtggcct ggcgcagctg 1260
 ccgtggttct taggggcttg gcccggggag cccatggggc ttgtgcctag aaagcctggg 1320
 ctccctactg gggcttagat gtgcagactt catgtctccc cagctccagc tctgttctct 1380
 ataggtcaag cctccacaat gccagaggcc cagggtcagc cccctccacg tccctcctag 1440
 atctacagct gcccttttga tgacagcgcc attgagtcct ctgggctggg ggggtcatgc 1500
 aggggtgagg cagctgcctg ccgccgggtac tcattgcctg gccaggcagg acacaggctg 1560
 gcgggcactg agagtggggc ccacgaaatc cattgtcagg ttaccaggat gaagaacca 1620
 ggctggtcgt ggagtgcagg gcggggcctg ccggaagaat tatgggcact gcagcaggag 1680

```

ggcagcctgg gccattagct cctgatgtca tcgatttggg tgaggggaca gggaagtcag 1740
aggaagctgg ccagtggtc tcacgcagac ttacagcagt ggagtgggtg ctgatttcctg 1800
gtacagctgc tcccactgag tctccaggga tctgtgggtc aggaccccct gcaaccccct 1860
cccagacccc tgtactgggt ggaggagagg acctagagga aaggtgctgg gcagataagc 1920
agctgaggga ggccctgggt ttagcttatc agtcttctgg gccctcctgc cccaggaagg 1980
gcagcgagga ccatgggtgt gccctgtca tcgttatcgt cctggccatg agcttgccag 2040
actgggaggg ccggagtcag ccaggcagac ggcagcacag catttgccgt ttggcagggtg 2100
gccttgggtg cttcccaaag gcaatcgctc caccgagaac aaaactcact tttttggggg 2160
gtgaagcacc ttgggttcatt tgtttagttc gtttaattcca gcagtctgtt tctaaggga 2220
acatgggtgc agccgggtcct gcgcctccca ccctcccacc aggtgcccag tgttcccaag 2280
ggccccgaat cccaacctta ttcaggcgtc agcatctctg caccctcaat gcctgttagg 2340
gaggatagtg aaggctgagc cctcctgggc ccatcaaaag ccagcagtga gagaacaccc 2400
ccatctctct gaggtgacct ttagggcag tccgtgctgt ctggctggcc tgggtgaggt 2460
gggcagggac caaggcctgg cgcctgggct tcgctggcct tgctctgctg gctgacttca 2520
tcctgatagt accttgattt tcctactgtg acttcccctt ctgtcgactt cctcaccaac 2580
tttaaaatc cgtattgaga gcagtttctt aagttacctc aaatcctatt cagaagaagg 2640
ttcttctctg aagttgggag ggccgaaaac aagtttagtc acagaagact actccatgtt 2700
tgagcttctg tttcaaggga agtgagtaac tgccggagga gccctgcccc tctgcagtgt 2760
gtggtgttgc cctgatactt ttcagattga ggtgttactt acatgtaata aaatgcacag 2820
acttaagtgt aaaaaaaaa aaaaaaaaa aaaaaaaaa 2860

```

<210> 56
 <211> 1559
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (1445)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1551)
 <223> n equals a,t,g, or c

```

<400> 56
atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagaccttct cctgcagctc 60
tacctacagc ggccggagct gcgggtgccc gtgcctgagg tcctactgca cagcgaaggg 120
gctgccagca gcagcgtctg caagctggac ggactcatcc accgcttcat cagctcctt 180
gcgacacca gcgactcccg ggcgttgagg aaccgagggg cggatgccag catggcctgc 240
cggaagctgg cgggtggcga cccgctgctg ctgctcaggc acctgcccac gatcgcgggc 300
ctcctgcacg gcgcaccca cctcaacttc caggagttcc ggcagcagaa ccacctgagc 360
tgcttcctgc acgtgctggg cctgctggag ctgctgcagc cgcacgtgtt ccgcagcgag 420
caccaggggg cgctgtggga ctgccttctg tccttcaccc gcctgctgct gaattacagg 480
aagtcctccc gccatctggc tgccttcac aacaagtttg tgcagttcat ccataagtac 540
attacctaca atgccccagc agccatctcc ttctgcaga agcacgcca cccgctccac 600
gacctgtcct tcgacaacag tgacctgtg atgctgaaat ccctccttgc agggctcagc 660
ctgccagca gggacgacag gaccgaccga ggcctggagc aagagggcca ggaggagagc 720
tcagccggct ccttgccctt ggtcagcgtc tccctgttca cccctctgac cgcggccgag 780
atggccccct acatgaaacg gctttcccgg ggccaaacgg tggaggatct gctggagggt 840
ctgagtgaca tagacgagat gtcccgccgg agaccggaga tcctgagctt cttctcgacc 900
aacctgcagc ggctgatgag ctccggccgg gagtggtgcc gcaacctcgc cttcagcctg 960
gccctgcgct ccatgcagaa cagccccagc attgcagccg ctttctctgc cagcttcag 1020
tactgcctgg gcagccagga ctttgagggt gtgcagacgg ccctccggaa cctgcctgag 1080
tacgctctcc tgtgccaaag gcacgcggct gtgctgctcc accgggcctt cctggtgggc 1140
atgtacggcc agatggaccc cagcgccgag atctccgagg ccctgaggat cctgcataag 1200

```

33

gagggcgtga	tgtgagcctg	tggcagccga	ccccctcca	agccccggcc	cgtcccgtcc	1260
ccggggatcc	tcgaggcaaa	gcccaggaag	cgtgggcgtt	gctggctctg	ccgaggaggt	1320
gagggcgccg	agccctgagg	ccaggcaggc	ccaggagcaa	tactccgagc	cctggggtgg	1380
ctccgggccc	gccgctggca	tcaggggccg	tccagcaagc	cctcattcac	cttctgggcc	1440
acagncctgc	gcggagcggc	ggatcccccc	gggcatggcc	tgggctgggt	ttgaatgaaa	1500
cgacctgaac	tgtcaaaaaa	aaaaaaaaaa	aaaccgrgg	gggggcccgg	naccctaatt	1559

<210> 57

<211> 2064

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (2001)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2024)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2049)

<223> n equals a,t,g, or c

<400> 57

atgggcgagg	ctgcggggcc	ccggcgcgca	cgcccgccacc	tctccccagc	cctggcgctgg	60
gcccagcccg	gcccaggcag	caatgggggt	cctgcagctg	ctggctcgtar	cggtgctggy	120
atccgaacac	cggttggtctg	gtgcagccga	ggtcttcggg	aattccagcg	arggtcttat	180
tgaattttct	gtggggaaat	ttagatactt	cgagctcaat	aggccctttc	cagaggaagc	240
tattttgcat	gatatttcaa	gcaatgtgac	ttttcttatt	ttccaaatac	actcacagta	300
tcagaataca	actgtttcct	tttctccgag	gcgtagatcc	cccacccatgt	gacgtctgga	360
cagaccagga	ctccagggtgg	aggttgcagt	atgatgtcta	tcagtatttt	ctgcctgaga	420
atgacctcac	tgaggagatg	ttgctgaagc	atctgcagag	gatggtcagt	gtgccccagg	480
tgaaggccag	tgctctcaag	gtggttaccc	taacagctaa	tgataagaca	agtgtttcct	540
tctctccct	cccgggacaa	ggtgtcatat	acaatgtcat	tgtttgggac	ccgtttctaa	600
atacatctgc	tgcttacatt	cctgtctaca	catagccttg	cagctttgag	gcaggagagg	660
gtagttgtgc	ttccctagga	agagtgtctt	ccaaagtgtt	cttcactctt	tttgccctgc	720
ttggtttctt	catttggttc	tttgacaca	gattctggaa	aacagaatta	ttcttcatag	780
gctttatcat	catgggatcc	ttcttttata	tactgattac	aagactgaca	cctatcaagt	840
atgatgtgaa	tctgattctg	acagctgtca	ctggaagcgt	cggtggaatg	ttcttggtag	900
ctgtgtggtg	gcgatttgga	atcctctcga	tctgcatgct	ctgtgttgga	ctagtgtctg	960
ggttcctcat	ctcgtcagtg	actttcttta	ctccactggg	aaacctaaag	atttttcatg	1020
atgatggtgt	attctgggtc	actttctctt	gcatagctat	cctcattcca	gtagttttca	1080
tgggctgcct	aagaatactg	aacatactga	cttgtggagt	cattggctcc	tattcggtgg	1140
ttttagccat	tgacagttac	tgggtccacaa	gcctttccta	catcactttg	aacgtactca	1200
agagagcgct	caacaaggat	ttccacagag	ctttcacaaa	tgtgcctttt	caaactaatg	1260
acttcattat	cctggcagta	tggggcatgc	tggctgtaag	tggaattacg	ttacagattc	1320
gaagagagag	aggacgaccg	ttcttccctc	cccaccata	caagttatgg	aagcaagaga	1380
gagagcgccg	agtgacaaac	attctggacc	ctagctacca	cattcctcca	ttgagagaga	1440
ggctctatgg	ccgattaacc	cagattaaag	ggctcttcca	gaaggagcag	ccagctggag	1500
agagaacgcc	tttgcttctg	tagatgccca	ggggcttggt	cagtgtgcct	cagctttgga	1560
gttcctgcct	ggagtgggtc	aacagtcctc	ggtgcaagtc	taataagaga	tcaggcatat	1620
atatctgttc	tttgcataat	attatgggtc	ccttattgat	atatggtaag	gggtgtactag	1680
gggattagga	tgattgtaag	agaatgagaa	agatgaccaa	aaggttggtg	gtagggaggc	1740

34

tttttcttat	ttccaaatac	ttgagaaatt	accttttggg	ttacaaatct	atgatcaact	1800
tattccatta	aatagatata	ttaaaaaat	taaaaaactga	ttcttctgca	gagcactggg	1860
gtttcttttt	ataaccctt	gaaacaagtc	tctcacctga	gcctgtctaa	actttcggag	1920
ggagtttatt	attgagtctt	tatctgtgac	agtatttggg	gatttaggga	tttgatactt	1980
aggcctttga	attttagaat	ncaaaaagg	gagccagcca	gacngggggg	ccaaccggga	2040
tccccaacng	ggaccagggg	ggtc				2064

<210> 58

<211> 1050

<212> DNA

<213> Homo sapiens

<400> 58

ccacgcgtcc	ggccagccag	tccgccgctc	cggagcccg	ctcgtctggg	cagcatggcg	60
gggtcgccgc	tgctctgggg	gccgcggg	gggggcgtcg	gccttttggg	gctgctgctg	120
ctcggcctgt	ttcggccgcc	ccccgcgctc	tgccgcgggc	cggtaaagga	gcccccgcc	180
ctaagcgcag	cgtctccgcc	cttggttaga	ctggcgctcc	tcgccgcttc	cggcggtcag	240
tgccccgagg	tgaggcggcg	ggggcggtgc	agacctggcg	cgggcgctgg	cgcactctgct	300
ggagccgaac	gtcaggagcg	ggcgcgggcc	gagcgcgaga	ggctgaggat	cagcaggcgc	360
gcgtcctggc	gcagctgctg	cgcgtctggg	gcgcccccg	caactctgat	ccggctctgg	420
gcctggacga	cgaccccgac	gcgcctgcag	cgcagctcgc	tcgcgctctg	ctccgcgccc	480
gccttgaccc	tgccgcccta	gcagccagc	ttgtccccgc	gcccgtcccc	gccgcggcgc	540
tccgaccccg	gcccccggtc	tacgacgacg	gccccgcggg	cccggatgct	gaggaggcag	600
gcgacgagac	acccgacgtg	gaccccgagc	tggtgaggtg	cttgctggga	cggattcttg	660
cgggaagcgc	ggactccgag	ggggtggcag	ccccgcgcgc	cctccgcctg	gccgcgcgac	720
acgatgtggg	ctctgagctg	ccccctgagg	gcgtgctggg	ggcgctgctg	cgtgtgaaac	780
gcctagagac	cccggcgccc	caggtgcctg	cacgcgcct	cttgccaccc	tgagcactgc	840
ccggatcccg	tgaccctg	gacccagaag	tgcccccgcc	atcccgccac	caggactgct	900
ccccgccagc	acgtccagag	caacttacc	cggccagcca	gcctctcac	ccgaggatcc	960
ctacccctg	gccccacaat	aaacatgatc	tgaagcagca	aaaaaaaaa	aaaaaaaaa	1020
aaaaaaaaa	aaaaaaaaa	aaaaaaaaa				1050

<210> 59

<211> 2533

<212> DNA

<213> Homo sapiens

<400> 59

ccacgcgtcc	gcctggcaac	ccctaatt	tggtatctct	caatgctatt	tgcttccatt	60
agcttgctcg	ttatgcttcc	cacttggtgg	attgtgtctt	cctggctggg	atggggagtg	120
attctatttg	tgatctggt	cataagagct	ttgagattat	ggaggacagc	caaactacaa	180
gtgaccttaa	aaaaatacac	cgttcatttg	gaagatatgg	ccacaaacag	ccgagctttt	240
actaacctcg	tgagaaaagc	tttacgtctc	attcaagaaa	ccgaagtgat	ttccagagga	300
tttacctg	tcagtgtg	ttgcccattt	aataaagctg	gacagcatcc	aagtcagcat	360
ctcatcggtc	ttcgaaagc	tgtctaccga	actctaagag	ccaacttcca	agcagcaagg	420
ctagctaccc	tatatatgct	gaaaaactac	ccccggaact	ctgagagtga	caatgtaacc	480
aactacatct	gtgtggtgcc	ttttaaagag	ctgggccttg	gacttagtga	agagcagatt	540
tcagaagagg	aagcacataa	ctttacagat	ggcttcagcc	tgccctgcat	gaagggtttg	600
ttccaaactct	gggtggcaca	gagttcagag	ttcttcagac	ggtagccct	attactttct	660
acagccaatt	cacctcctgg	gcccttactt	actccagcac	ttctgcctca	tcgtatctta	720
tctgatgtga	ctcaaggctt	acctcatgct	cattctgcct	gtttggaaga	gcttaagcgc	780
agctatgagt	tctatcggtg	ctttgaaact	cagcaccagt	cagtaccgca	gtgtttatcc	840
aaaactcaac	agaagtcaag	agaactgaat	aatgttcaca	cagcagtgcg	tagccttgag	900
ctccatctga	aagcattact	gaatgaggtg	ataattcttg	aagatgaact	tgaagagctt	960
gtttgtacta	aagaaacaca	agaactagtg	tcagaggctt	atccccatcc	agaacagaaa	1020
ttaaagttga	ttcagcccca	cgttcaagca	agcaacaatt	gctgggaaga	ggccatttct	1080

35

caggctcgaca	aactgctacg	aagaaataca	gataaaaaag	gcaagcctga	aatagcatgt	1140
gaaaaccacac	attgtacagt	aagtaccttt	gaagcagcct	actctacaca	ttgcagacaa	1200
agatccaatc	ccagaggagc	aggaattaga	agcttatgta	gatgatatag	atattgatag	1260
tgatttcaga	aaggatgatt	tttattactt	gtctcaagaa	gacaaagaga	gacagaagcg	1320
tgagcatgaa	gaatccaaga	gggtgctcca	agaattaaaa	tctgtgctgg	gatttaaagc	1380
ttcagaggca	gaaaggcaga	agtggagca	acttctat	agtgatcatg	tgtttcttca	1440
tatagcttta	aaattatgct	attgacatta	tgggaaagat	ttatcaatga	gagaaatgtg	1500
tctctttttc	agccgtgttg	aaatccttgt	ctcctgtaga	cccagtggaa	cccataagta	1560
attcagaacc	atcaatgaat	tcagatatgg	gaaaagtcag	taaaaatgat	actgaagagg	1620
aaagtaataa	atcccgcaca	acagacaatg	aaataagtag	gactgagtat	ttatgtgaaa	1680
actctctaga	aggtaaaaa	aaagataatt	cttcaaatga	agtcttcccc	caaggagcag	1740
aagaaagaat	gtgttaccaa	tgtgagagtg	aagatgaacc	acaagcagat	ggaagtgggtc	1800
tgaccactgc	ccctccaact	cccagggaact	cattacagcc	ctccattaag	cagaggctgg	1860
cacggctaca	gctgtcacca	gattttacct	tcactgctgg	ccttgctgca	gaagtggctg	1920
ctagatctct	ctcctttacc	accatgcagg	aacagacttt	tggtgatgag	gaggaagaac	1980
aaataataga	agaaaaataa	aatgagatag	aagaaaagta	agaaccaaga	ttcatatgaa	2040
gtgatattag	attgttcctt	ttacaaaagt	gtttagcttc	aagactggaa	agggaatatg	2100
agtgtaatgt	tactatatat	aaagctaaga	tgtggattta	caggaagaac	cctggtttga	2160
ataactgatc	tgaaattagt	agttacctgt	aaatggcaga	tcttttagga	aaataagaga	2220
aaggtaaggg	ctcttttgaa	taaactgctg	ttttatttgt	ggcacaaactg	atcaatcttg	2280
gaaattcttt	aagtattttt	aataagaaat	gaattatcat	ttcttgccag	aatttgctac	2340
cttaaggtga	ttgggaaaat	tctgttgcaa	gaacattaac	atttagtatg	actccttttt	2400
actgtattct	tgagtgtaat	aactgcagct	attatgttaa	taacaagttg	tttgtatttt	2460
atttttgttt	ataccagtct	taaagatcca	ggttctgaa	aaaaaaatta	attgatacaa	2520
aaaaaaaaaa	aaa					2533

<210> 60

<211> 899

<212> DNA

<213> Homo sapiens

<400> 60

ggcagatttc	ccggcacctt	cgtgggcacc	acagagcccg	cctccccacc	cctgagcagc	60
acctcaccca	ccactgctgc	ggccactatg	cctgtggtgc	cctctgtggc	cagcctggcc	120
cctccggggg	aggcctcgct	ctgcctggaa	gaggtggccc	ccccgccag	tgggacccgc	180
aaagctcggg	tgctctatga	ctacgaggca	gccgacagca	gtgagctggc	cctgctggct	240
gatgagctca	tcactgtcta	cagcctgcct	ggcatggacc	ctgactggct	cattggcgag	300
agaggcaaca	agaagggcaa	ggtccctgtc	acctacttgg	aactgctcag	ctaggcaggt	360
gccccatcc	ccccgcatt	ctggcctagg	caggagagga	tgggcgact	gccacttaac	420
ttgtttgttg	tgacacag	tggtcagagt	ggggagaatt	caccccatc	tgctccctgc	480
cctagtcacc	tagctgtgag	ggtgcctgag	gctgaatggc	tccaccctcc	cccagccctg	540
cttctgacct	gtggctctgg	agccccctgc	cctgcctgca	tccccgagca	ccccaccctc	600
caggctccac	taaggaggga	ggggctgtct	gcagcagctg	cactcagcac	ctaggccagg	660
gtggggccgc	gcagatggg	ctcaggaagc	cccaggtgca	ctcagcgaga	gccctgcctt	720
tcagttgcca	aaagctgcat	caggggaatg	cggcaaggca	cacagggctc	tgccagcccc	780
tggggactgg	gcgctgcccc	tgggagggga	gagcctggcc	agggctggtg	ttgggcccgg	840
agcagcatct	tccggtgcta	tcctccctcc	ccaccctca	cagctcaagc	caagtcag	899

<210> 61

<211> 1079

<212> DNA

<213> Homo sapiens

<400> 61

tcgacccacg	cgtccgggtt	tcaccacgtt	ggccaggctg	gtctcaaact	cctgacttca	60
gttgccctcc	caaagtgtcg	ggattaaagg	catgagccac	tgccgccggc	ctacccttct	120

aactctactt	ctagcttctt	gcttctgggc	tgctgctata	ccaaacagga	atgtaatact	180
ttctgtcagc	ttcaggcctt	tgacatgca	gttactttt	tctatcttgg	tttttattct	240
taggatttta	attctcctaa	gaagctttct	ctgaccagcc	taaaacttac	gtaagccctg	300
ggttaggtgc	tatgcttatg	tcctcccata	gcattttgca	tttgcatgtg	ttgtaactct	360
taatgtacag	catcatgatt	gcctatttta	actttcctgt	ttgttacagt	agactttaat	420
ctctttaagg	acaggaactg	tgtcttggtt	agaatcccca	gagcttattt	agtacaatgg	480
ctatgcttat	aatttaagta	tttattgaac	aaatgaaatt	ttcctaagcc	ctaaaacctt	540
gcaagatggt	ttagtgaggg	aaactggcct	cgggtggagt	gaataactag	cacgaggtca	600
ctcacctaaa	aagtgggtgag	gagggattaa	aatctaaatc	tgtttagctg	taaagattgg	660
gcttttttyc	ttgctgctgc	acatgactgc	ytctctctcat	gttgccctgta	cacatccctg	720
tcaagtgttc	aaacagcccg	tgccctaaca	ccccatccat	agcttctgag	gaaagtgtg	780
tcatctttgg	acagctctga	gagctgaagc	gagctcttgc	agaataattt	cccatctatt	840
ggctcttaatt	tatgctttgg	agaatataac	ttattttcaa	aaaacaaatg	attcagaatt	900
tgctcatctcc	ttaaggtccg	tttattagtt	tatttcattc	cttcattcac	tgataacccat	960
ttactgagca	ccagcctggg	caacatggtg	agaacccatc	tctaccaatt	taaaaaaaaa	1020
aaaaaagggc	ggccgctcta	gaggatccaa	gcttacgtac	gcgtgcatgc	gacgtcata	1079

<210> 62

<211> 1928

<212> DNA

<213> Homo sapiens

<400> 62

ggcacgagag	taggtctgcc	ggcgatggag	tggtgggcta	gctcgccgct	tcggctctgg	60
ctgctgttgt	tcctcctgcc	ctcagcgag	ggccgccaga	aggagtcagg	ttcaaatgg	120
aaagtattta	ttgaccaa	taacaggtct	ttggagaatt	acgaaccatg	ttcaagtcaa	180
aactgcagct	gctaccatgg	tgctcatagaa	gaggatctaa	ctcctttccg	aggagggcatc	240
tccaggaaga	tgatggcaga	ggtagtcaga	cgaagctag	ggaccacta	tcagatcact	300
agaacagac	tgtaccggga	aaatgactgc	atgttcccct	caagggtgtag	tggtgttgag	360
cactttattt	tggaagtgat	cgggcgtctc	cctgacatgg	agatgggtgat	caatgtacga	420
gattatcctc	agggttcctaa	atggatggag	cctgccatcc	cagtcttctc	cttcagtaag	480
acatcagagt	accatgatata	catgtatcct	gcttggacat	tttgggaagg	gggacctgct	540
gtttggccaa	tttatectac	aggctcttga	cgttgggacc	tcttcagaga	agatctggta	600
aggctcagag	cacagtggcc	atggaaaaag	aaaaactcta	cagcatattt	ccgaggatca	660
aggacaagtc	cagaacgaga	tcctctcatt	cttctgtctc	ggaaaaaacc	aaaacttggt	720
gatgcagaat	acaccaaaaa	ccaggcctgg	aaatctatga	aagatacctt	aggaaagcca	780
gctgctaagg	atgtccatct	tgtggatcac	tgcaaataca	agtatctggt	taattttcga	840
ggcgtagctg	caagtttccg	gtttaaacac	ctcttctgt	gtggctcact	tgttttccat	900
gttggtgatg	agtggctaga	attcttctat	ccacagctga	agccatgggt	tcactatata	960
ccagtcaaaa	cagatctctc	caatgtccaa	gagctgttac	aatttgtaaa	agcaaatgat	1020
gatgtagctc	aagagattgc	tgaaggggga	agccagttaa	ttaggaacca	tttgcatgatg	1080
gatgacatca	cctgttactg	ggagaacctc	ttgagtgaat	actctaaatt	cctgtcttat	1140
aatgtaacga	gaaggaaagg	ttatgatcaa	attattccca	aaatgttgaa	aactgaacta	1200
tagtagtcat	cataggacca	tagtcctctt	tgtggcaaca	gatctcagat	atcctacgggt	1260
gagaagctta	ccataagctt	ggcacctata	ccttgaatat	ctgctatcaa	gccaaatacc	1320
tggttttctt	tatcatgctg	cacccagagc	aactcttgag	aaagatttaa	aatgtgtcta	1380
atacactgat	atgaagcagt	tcaacttttt	ggatgaataa	ggaccagaaa	tcgtgagatg	1440
tggaatttga	acccaactct	acctttcatt	ttcttaagac	caatcacagc	ttgtgcctca	1500
gatcatccac	ctgtgtgagt	ccatcactgt	gaaattgact	gtgtccatgt	gatgatgccc	1560
tttgtcccat	tatttgagag	agaaaattcg	tcatttggaa	gtagtacaac	tcattgctgg	1620
aattgtgaaa	ttattcaagg	cgtgatctct	gtcactttat	tttaatgtag	gaaacccctat	1680
ggggtttatg	aaaaatactt	ggggatcatt	ctctgaatgg	tctaaggaag	cggtagccat	1740
gccatgcaat	gatgtaggag	ttctcttttg	taaaaccata	aactctgtta	ctcaggaggt	1800
ttctataatg	ccacatagaa	agaggccaat	tgcatgagta	attattgcaa	ttggatttca	1860
ggttcccttt	ttgtgccttc	atgccctact	tcttaatgcc	tctctaaagc	caaaaaaaaa	1920
aaaaaaaa						1928

<210> 63
<211> 781
<212> DNA
<213> Homo sapiens

<400> 63
ggcacgagat tttcagcctt tttggactgg tttctccaca tcttcgtgga tttatctaac 60
tttggctctt gatgttggtg accttcagat tgggtctctg agtgaacatc ctttttgggtg 120
atgttgatac tattcctttc tggttggttg tttgttttcc ttctaacagt cagggccctc 180
tgctgcaggt ctgctggagt ttggttgagg tccactccag accctgtttg tctgggtttt 240
gccagaggag gctgcagaat agcaatgatt gctgcctgtt ttccctctgg aagctttgtc 300
ccagaggggc acccaccaga tgccagccag agctctcctg tatgaggtgt ctgttgcccc 360
atacttggag gtgccttcca gtcaggatac acaggtgtca ggtaccactc tgaggaggca 420
ctctgtcccc tatcagagct cgaacactgt gctggggagat ccactgttct cttcagagct 480
gtcagacagg gacgtttaag tctgctgaag ctatgccac agctgcccct tccccagat 540
gctctgtccc agggagaagg gagttttatc tataagtctc tgactggggc tgctgccttt 600
tcttcagaga tgccctgccc caagacgggg actctagaga ggcagtctgg ctgcagtggc 660
cttgctgaac tgggtggggc ttcacccagt tggaccttcc ctgagccttt ttttttcccc 720
tgtgagggtg aaaatgccta atcaagcctc agcaatggtg gatgccttc cccccaccaa 780
g 781

<210> 64
<211> 1194
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1172)
<223> n equals a,t,g, or c

<400> 64
ggcacgagaa gacatggagt cttaagtgtg atcagtggga gggggctgga atcatttaga 60
ggcatcttca ttcacaaaac caggagctga tactggctgt cagccaggac ttcaactgac 120
ctatgtagaa cctgtccatg tggcccttcc ttgcagtctc cccatttggg ctggtttggg 180
cttcatcaca gtccggcagc ttacttctaa gggcaagcat tccacgacaa cacagcagaa 240
gggcatggca tttttacagt gaagtttggc aatctcatag cgtcgcttct gtcctacttt 300
atttattggt cagggcaatc acaaagatgt gcataggctc aaagaaaaga gacataaccc 360
cgaccacgag atggaagaag tgacacggtc atgttatgag aggagtgtgt gggatgggag 420
atagggctgt ggccacctgc agaaaacagc atctgctata ggctgtcatg gaagcgcagg 480
atggggattt agcctacctg aggggtcagt cagcaaaggc ctctggggagg aagtggatgc 540
ttcggctgag gatgtgaagg gctaaaagga gaatgaggaa gagtttcagg gagaggaatc 600
aatgaaacga gtccagagac gctggtgagt tggatggtt gcttcagtat gatgacaata 660
cagaggggca aggagactgg tgcaggagaa gagagaaggt gccatgtgct ctgggtcgtg 720
tcttctatgc cagactccct tagaagagga gcagcctcca gtcagcgtgt tcccaggaaac 780
acggaggcta gacaggacaa tggcagccaa tccctgctcc caaactgggt acagtgggga 840
aaagctgcat ggtctagatc caccctgctc cctggcccca gtatagaaga tcaaattcaa 900
tctgccaat cttatccaga taaagtaaag gaagactgga aaaaagaact aatccacggc 960
tccatctgcc catgactttc tctgctgatg cgggaggcag ctatggataa agagacggca 1020
cacggcatgt cccgacgctg tggaggtggg gagacccgc aagtccacag gaaaagagtt 1080
aagttgctgc cactgggca tccgctattc tctgctcttc tgcctcatcc tcaattcaga 1140
ccatgatgga gctgattgtc tcccatttta tnccttggat tgaatggtct cgag 1194

<210> 65
<211> 1677

<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1012)
<223> n equals a,t,g, or c

<400> 65
ggtgcagtgg tgccatcaca gttcactgca gccttgacct cccgggctca agcaatcctc 60
ccacctcagc cacttgagta gctgagacct cagatatgtg ccatcacacc cagctgattt 120
tttaaaatta attttttgta gagatagggt ctcatatgtt gccatgctg gtctcaaaact 180
actgggttca aatgatcctc ctgcctcagc cttccaaagt actgggatta caggcatgag 240
ccaccatgcc gggctgggag gcggaatttt gttcagtcta aagataagct ttttcatagc 300
tctggctgta gtgggagggg gcagaggagt gaattgattgt cagttgggag ggtgcagagt 360
gggctcctgc cctagggtgr aggttagggg ggcttaggtg asmcamcaca gaggccctgt 420
tcagcccccac gtccccctccc tgtgtctccc cctcctctct cctcctcctgc aggcgtggga 480
ggtatcatca ttcagcagat ttcaccagag gcagtggagg aggcaggtac ctgagccaga 540
attcagaatg tcttattctc cacttgactc tgccactaac ttgttggtgca actttgggcc 600
tttccccagg ccttcatttt cttttctttt ctttttcttt yttttttttt gaggcggagt 660
ctcgctatgt tgcccaggct ggagtgaggt ggcgcagcat catctcggct cactgcaagc 720
tccaccttct gagttcacgc cattctactg cctcagcctc ccgagtagcc gggactgcag 780
gcrccacca ccacgcccg cttatttttt gtatttttag tagagacagg gtttcaccac 840
gttagccaag atggtctcga tctcctgacc tcgtgatcca cccgcctggg cttcccaaag 900
tgctgggatt acaggcgtga gccactgcgc ccggccattt tcttaaatat ctaataaaaa 960
atatatagca aatgcagttt ttaactacg acaatatgac cagcgaagag antattatct 1020
tccaagactg ctggtccaag gaaaagtcag taataaagtg gaagcattgt agcttatgga 1080
atgactggtt asatttggga gaagccttag caataatcta gaatctgcat agataatata 1140
tctgaggatt gggctttgtg gtttacaag catttttttt tctctctttg atcccagccg 1200
tttgtctgga ctgatacaaa gcatttttat tagtttctct tattcaatcc tcacaccacc 1260
tcaaatttac agaggatatg gatctgggta atttgtatga ctatgtaacc tcatgtcagt 1320
ccacagcact gcctggaggt gggtagaggt ggtcctgggc tggaaatcca gccccagtg 1380
gaccttgagc aagttaacttt agctgtctgc acctaaattt cctcactggc aaaacaggaa 1440
tactggtggt tcacacctgc aattccagca ctttgggagg ctgaggtggg aggattgctt 1500
gagtcagaa gttcaaaaac agactgggca acatagcaag accatctcta caaaaattaa 1560
ataaataaaa catttacaag ggttggtgtg aagattaaat gagatcactc acgaaaaagc 1620
tcagcagacc ctgatgtgca gtaggtgctc aataaatgtt agccagcaaa aaaaaag 1677

<210> 66
<211> 1237
<212> DNA
<213> Homo sapiens

<400> 66
agcaaaccca ggaaggtgtg gcgtccccgc ttcgcgccaa gatggtgctg gtgctgcgcc 60
atcctttgtg tgcccgggaa agggcggtcc gggagccggg tcgggggctc ctgactcgca 120
ctgggcagca tgacggtgcy ccggctgtca ctgctgtgcc gggacctctg ggcgctgtgg 180
ctgctgctga aggcggcgcc agtgctgtgg gcgcgggcgg gtccctgcct ccccggaagg 240
tgttggtggg cgacatgcgg ggaacccggg cgggggtgga cgttctgggc ccagccctgt 300
cctcagaagc tgctggggca gaagcccggg gctgggggat gccggggatg ggtgttgggg 360
tggtgtcctc cgagaccaga ggagccctgt tccttgccag ggaaggtgtg cacgggcctt 420
gcccgatgga tggtttaggg ccatggccct ggggtccctg gtgagcagtg gggccgcctc 480
tgcccttggc ctgtgaggga ctgtctgtgc tgggccaga aggttgggat cacctttcca 540
ctggctcctt tgttcgaggt ttttcataga caggctatgt ggacaaatga gggcagcgcc 600
cacgtctggc tgggtggagg gctgcggctc ctcttggag gggacgcctg gccactgctg 660
tccccacaat ggggccaccg gtggtgcaag gcgtgacaag ctgccctctc taggtaagca 720
ggacttggga ggccctggc caagcctgtg gaccggcctg ggcggcctct gtggtctcag 780

39

gtttgggtgt	gtttgggtctg	gtcagggctc	aggggctgct	ggccacact	ggccccatcc	840
tgacaattgg	agctttgggg	caaggtccct	ggagaagggg	tcacgtcggg	aggaacacgc	900
ctgggttttg	ttgatgcttt	tctaagaatg	gagtactcgt	tttcaagaga	tttgctctaa	960
ttatattttc	cagcgggtac	ttatgccaag	tattgatgaa	taattcataa	aataagcatc	1020
tttgtgaatt	ttagtgaatc	agaccttaac	tatcaacggc	aatgaatgaa	catctaaagt	1080
ttccaatttt	aaagtaaaga	actggctggg	tacagcagtt	cacgcctgta	atcccagcac	1140
tttgggaggg	caaggctaga	ggatcgcttg	agcccaggag	tttgagatca	gcctggggcaa	1200
cataccaaga	cctcatctgt	taaaaaaaaa	aaaaaaa			1237

<210> 67
 <211> 1934
 <212> DNA
 <213> Homo sapiens

<400> 67

ccacgcgtcc	ggggcggtcc	tggtcgtgag	aggggagccc	caggggagct	ggggcagcat	60
gactgggggtg	ataaatggcc	ggaaatttgg	cgtggccaca	ctcaacacca	gcgtgatgca	120
ggaggcacac	tccgggggtca	gcagcatcca	cagcagcatc	cgccatgtcc	cagcaaacgt	180
ggggcctctg	atgcgggtgc	tcgtgggtcac	catcgccccc	atctactggg	ccctggccag	240
agagagtgagg	gaagccctga	atggccactc	tctgactggg	ggcaagtctc	ggcaggagtc	300
acacgtggag	tttgctacag	gggagctgct	cacgatgacc	cagtggcccc	gggtctggat	360
cccgatggcc	tcctgctcct	cgacgtggtg	gtcaatggcg	ttgtccccgg	acagcctggc	420
tgacgcagat	cttcaagtgc	aggactttga	ggagcactac	gtgcaaacag	ggcctggcca	480
gctgttcgtg	ggctccacac	agcgcttctt	ccagggcggc	ctccccctgt	tcctacgctg	540
caaccacagc	atccagtaca	acgcggcccg	ggggccccag	ccccagctgg	tgcagcacct	600
gcgggcctca	gctatcagct	cggcctttga	tccagaggcc	gaggccctgc	gcttccagct	660
cgctacagcc	ctgcaggcgg	aggagaacga	ggtcggctgc	cccaggggct	ttgagctgga	720
ctcccaggga	gcgtttttgtg	tggatgtgga	cgagtgtgcg	tgggatgtc	acctctgccg	780
agagggacag	cgctgtgtga	acctgctcgg	gtcctaccgc	tgcctccccg	actgtggggc	840
tggcttccgg	gtggctgatg	gggccggctg	tgaagatgtg	gacgaatgcc	tggagggggt	900
ggacgactgt	cactacaacc	agctctgcga	gaacacccca	ggcggtcacc	gctgcagctg	960
ccccagggggt	taccggatgc	agggccccag	cctgccctgc	ctagatgtca	atgagtgcct	1020
gcagctgccc	aaggcctgcg	cctaccagtg	ccacaacctc	cagggcagct	accgctgcct	1080
gtgcccccca	ggccagaccc	tccttcgcga	cggcaaggcc	tgcacctcac	tggagcggaa	1140
tggacaaaat	gtgaccaccg	tcagccaccg	aggccctcta	ttgccctggc	tgcggccctg	1200
ggcctcgatc	cccggtagct	cctaccacgc	ctgggtctct	ctccgtccgg	gtccccatggc	1260
cctgagcagt	gtgggcccgg	cctgggtccc	tcctggtttc	atcaggcaga	acggagtctg	1320
cacagacctt	gacgagtgcc	gcgtgaggaa	cctgtgtcag	cacgcctgcc	gcaacactga	1380
gggcagctac	cagtgcctgt	gccccgcccg	ctaccgtctg	ctccccagcg	ggaagaactg	1440
ccaggacatc	aacgagtgcg	aggaggagag	catcgagtgt	ggaccgggcc	agatgtgctt	1500
caacacccgt	ggcagctacc	agtgtgtgga	cacaccctgt	cctgccacct	accggcaggg	1560
ccccagccct	gggacgtgct	tccggcgctg	ctcgcaggac	tgcggcacgg	gcggcccttc	1620
tacgctgcag	taccggctgc	tgcgctgcgc	cctggggcgtg	cgcgcccacc	acgacgtggc	1680
ccgcctcacc	gccttctccg	aggtcggcgt	ccccgccaac	cgcaccgagc	tcagcatgct	1740
ggagccccgac	ccccgcagcc	ccttcgcgct	gcgtccgctg	cgcgcggggc	ttggcgcggt	1800
ctacacccgt	cgcgcgctca	cccgcgccgg	cctctaccgg	ctcaccgtgc	gtgctgcggc	1860
accgcgccac	caaagcgtct	tcgtcttgct	catcgccgtg	tccccctacc	cctactaaac	1920
gggagaggggc	attg					1934

<210> 68
 <211> 3300
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE

<222> (1)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (3)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (15)
 <223> n equals a,t,g, or c

<400> 68
 ncngcagccg gacgnccgag cgcagcgagt cagtgcgcga ggaagcggaa gagcgcccaa 60
 tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg cagcacaggt 120
 ttcccgcactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag ctccactcatt 180
 aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga attgtgagcg 240
 gataacaatt tcacacagga aacagctatg accatgatta cgccaagctc gaaattaacc 300
 ctactaaag ggaacaaaag ctggagctcc accgcggtgg cggccgctct agaactagt 360
 gatcccccg gctgcaggaa ttcggcacga gaacacatct taagggaacc aagtctcaag 420
 agaaatcaag taattatgaa tgaacagctc taaaaaagag agagagaata ttttcttaaa 480
 tcaacttagt tgctgttatg accaaagaac agatgttgtg gtgttcaccc cagagaagca 540
 agagattttc ccttaaacct cagcttataa tgaatggaa gaaatgacag ggagagagtt 600
 tttctctcgt ttcccagaac tctatccttt tcttctcaaa cagttggaaa ctgtagccaa 660
 tacagttagc agtgatatgg gagaaccaa tcgtcatcca agcatgtttc tcttactttt 720
 ggtgttgagg agactctacg cttccccgat ggatgggtact tcttctgctc tcagcatggg 780
 accttttgtt cccttcatta tgagggtgtg tcactcacct gtctaccact ccgctgaat 840
 ggcagctcgt gccttggtcc catttggtat gatagatcac attcctaata ccattcgaac 900
 tctgttgtcc acactcccca gctgcactga ccagtgttc cggcaaaaacc acattcatgg 960
 gacacttttc cagggttttc atttgttgca agcctactca gactccaaac acggaacgaa 1020
 ttcagacttc cagcacgagc tgactgacat cactgtttgt accaaagcca aactctggct 1080
 ggccaagagg caaaatccat gtttggtgac cagagctgta tatattgata tttctctcct 1140
 attgacttgc tgccctcaaca gatctgcaa ggacaaccag ccagtctctg agagtcttgg 1200
 cttctgggag gaagtcagag ggattatctc aggatcagag ctgataacgg gattcccttg 1260
 ggccttcaag gtgccaggcc tgccccagta cctccagagc ctcaccagac tagccattgc 1320
 tgcagtgtgg gccgcggcag ccaagagtgg agagcgggag acgaatgtcc ccactctctt 1380
 ctctcagctg ttagaatctg ccttccctga agtgcgctca ctaacactgg aagccctctt 1440
 ggaaaagtcc ttagcagcag cctctggact tggagagaag ggcgtgccac cctgtctgtg 1500
 caacatggga gagaagtctt tattgttggc catgaaggaa aatcaccag aatgctctg 1560
 caagatactg aaaattctcc actgcatgga ccctgggtgag tggcttcccc agacggagca 1620
 ctgtgtccat ctgaccccaa aggagtctt gatctggacg atggatattg cttccaatga 1680
 aagatctgaa attcagagt tagctctgag acttgcttcc aaagtcattt cccaccacat 1740
 gcagacatgt gtggagaaca gggaattgat agctgctgag ctgaagcagt ggtttcagct 1800
 ggtcatcttg tcatgtgaag accatcttcc tacagagtct aggttgccg tcgttgaagt 1860
 cctcaccagt actacaccac ttttctcac caaccccat cctattcttg agttgcagga 1920
 tacacttgct ctctggaagt gtgtccttac ccttctgacg agtgaggagc aagctgttag 1980
 agatgcagcc acggaaaccg tgacaactgc catgtcacaa gaaaatacct gccagtcac 2040
 agagtttgcc ttctgccagg tggatgcctc catcgctctg gccctggccc tggccgtcct 2100
 gtgtgatctg ctccagcagt gggaccagtt ggcccttggc ctgcccattc tgcgtggatg 2160
 gctgttgga gagagtgat acctcgtggc ctgtgtggag agcatgcac aggtggaaga 2220
 agactacctg ttgaaaaag cagaagtcaa cttttgggcc gagaccctga tctttgtgaa 2280
 atacctctgc aagcacctct tctgtctcct ctcaaagtcc ggctggcgtc cccaagccc 2340
 tgagatgctc tgtcaccttc aaaggatggg gtcagagcag tgccacctcc tgcctcagtt 2400
 cttcagagag ctccaccag ctgctgagtt tgtgaagaca gtggagttca caagactacg 2460
 cattcaagag gaaaggactt tggcttgctt gagcgtgctg gccttttttg aaggaaagga 2520
 aggggaagac accctagttc tcagtgtttg ggactcttat gcagaatcga ggcagttaac 2580
 tcttccaaga acagaagcgg catgttgaag aaaatctggg ggattgggat gggggtatgt 2640

41

gtggattttt	cctccactaa	atctgcagga	aacatgttga	acataaatcc	aaaaatttta	2700
tccccaaaaa	aaaaaaaaaa	aaactcgagg	gggggcccgg	tacccaattc	gccctatagt	2760
gagtcgtatt	acaattcact	ggccgtcggt	ttacaacgtc	gtgactggga	aaacctgggc	2820
gttaccacaac	ttaatcgctt	tcgagcacat	ccccctttcg	ccagctggcg	taatagcgaa	2880
gaggcccgcga	ccgatcgccc	ttcccacag	ttgcgcagcc	tgaatggcga	atggcaaat	2940
gtaagcggtta	atattttgtt	aaaattcgcg	ttaaattttt	gttaaatcag	ctcatttttt	3000
aaaccaatagg	ccgaaatcgg	caaaatccct	tataaatcaa	aagaatagac	cgagataggg	3060
ttgagtgttg	ttccagtttg	gaacaagagt	ccactattaa	agaacgtgga	ctccaacgtc	3120
aaagggcgaa	aaaccgtcta	tcagggcgat	ggccactac	gtgaaccatc	accctaataca	3180
agtttttttg	ggtcgaggtg	ccgtaaagca	ctaaatcgga	accctaagg	gagcccccga	3240
tttagagctt	gacggggaaa	gccggcgaa	gtggcgagaa	aggaaggga	gctgtctctt	3300

<210> 69

<211> 1797

<212> DNA

<213> Homo sapiens

<400> 69

ggtcgacggt	atcgataagc	ttgatatcga	attcctgcaa	cagttcttgg	aaaccactc	60
gagagggcca	cgccctccatt	caccaggcca	cgcatcacia	gaggcaacac	caggagccaa	120
catgagctcg	gggactgaac	tgctgtggcc	cggagcagcg	ctgctgtgtc	tggtgggggt	180
ggcagccagt	ctgtgtgtgc	gctgtctcac	cccaggtgca	aagaggtcag	agaaaaatcta	240
ccagcagaga	agtctgctg	aggaccaaca	gagctttacg	gggtcccggg	cctactcctt	300
ggtcgggag	gcattggccag	gaccctgtgc	ggacatggca	cccacaagga	aggacaagct	360
gttgcaattc	taccccagcc	tgaggagatcc	agcatcttcc	aggtaccaga	acttcagcaa	420
aggaagcaga	cacgggtcgg	aggaagccta	catagacccc	attgccatgg	agtattacaa	480
ctgggggagg	ttctcgaagc	ccccagaaga	tgatgatgcc	aattcctacg	agaatgtgct	540
catttgcaag	cagaaaaacca	cagagacagg	tgcccagcag	gagggcatag	gtggcctctg	600
cagaggggac	ctcagcctgt	cactggccct	gaagactggc	cccacttctg	gtctctgtcc	660
ctctgcctcc	ccggaagaag	atgaagggaat	ctgaggatta	tcagaacttc	agcattccat	720
ccattcagtg	gcgcgagtc	aggaaggtca	tggggcaact	ccagagaaga	aagcatcccc	780
tgccccggtg	ggaagcccag	acgaggagga	cggggaaaccg	gattacgtga	atggggaggt	840
ggcagccaca	gaagcctagg	gcagaccaag	aagaaaggag	ccaaggcaaa	gagggaccac	900
tgtgctcatg	gaccatcgc	tgccctccaa	ggaccatttc	ccagagctac	tcaactttta	960
agcccctgcc	atggttgctc	ctggaaggag	aaccagccac	cctgaggacc	acctggccat	1020
gcgtgcacag	cctgggaaaa	gacagttact	cacgggagct	gcaggccccc	tcaccaagcc	1080
ctctcccagc	ccaggctttg	tggggagagg	acctggtacc	aagggttaacc	cggctcctgg	1140
tatggacgga	tgccgaggat	ttaggataag	ctgtcaccca	gtccccataa	caaaaccact	1200
gtccaacact	ggatatctgt	ttcttttggt	ctatgaattt	ggattcctaa	ttgctattgt	1260
tggttgctgg	ggtttttaaa	gattgataag	ctgtacagct	taacttatag	agggggagcc	1320
atatttaaca	ttctggattt	cagagtagag	atttctgtgt	tgctcctag	aaagcattac	1380
atgtagttaa	tttcagcatc	cttggtgggt	ggggccctgg	ctctcttccc	ctttggtggg	1440
acctcccctt	tccttgggct	tcagttcact	caggaagaaa	tgaggctgtc	gccatcttta	1500
tgtgcttcca	gtggaaatgt	cacttgctac	agacaatagt	gcatgagagt	ctagagaagt	1560
agtgaccaga	acagggcaga	gtaggtcccc	tccatggccc	tgaatcctcc	tctgtcccag	1620
ggctggcctc	tgacagagctg	attaaacagt	gttgtagctg	tctcatggga	agagctgggg	1680
cccagaggga	ccttgagtca	gaaatgttgc	cagaaaaagt	atctcctcca	accaaaccat	1740
ctcaataaaa	ccattttagt	tgaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaa	1797

<210> 70

<211> 1373

<212> DNA

<213> Homo sapiens

<400> 70

ggcacgaggg	ctgacggcgc	ttttgtctcc	ggtgagtttt	gtggcgggaa	gcttctgcgc	60
------------	------------	------------	------------	------------	------------	----

tggtgcttag	taaccgactt	tcttccggac	tcttgcacga	cctgctccta	cagccggcga	120
tccactcccg	gctgttcccc	cggagggtcca	gaggcctttc	agaaggagaa	ggcagctctg	180
tttctctgca	gaggagtagg	gtccttttcag	ccatgaagca	tgtgttgaac	ctctacctgt	240
taggtgtggt	actgacccta	ctctccatct	tcgttagagt	gatggagtcc	ctagagggt	300
tactagagag	cccatcgctt	gggacctcct	ggaccaccag	aagccaacta	gccaacacag	360
agcccaccaa	gggcttccca	gacctccat	ccagaagcat	gtgataagac	ctccttccat	420
actggccata	ttttggaaca	ctgacctaga	catgtccaga	tgggagtccc	attcctagca	480
gacaagctga	gcaccgttgt	aaccagagaa	ctattactag	gccttgaaaa	acctgtctaa	540
ctggatgctc	attgcctggg	caaggcctgt	ttaggcctgt	tgcgggtggc	catgcctgta	600
atcctagcac	tttgggaggc	tgatgtgggt	ggatcacctg	aggtcaggag	ttcagaccag	660
cctcgccaac	atggcgaaac	cccatctcta	ctaaaaatac	aaaagttagc	tgggtgtggt	720
ggcagaggcc	tgtaatccca	gtccttggg	aggctgaggc	gggagaattg	cttgaacccg	780
gggacggagg	ttgcagttag	ccgagatcgc	actgctgtac	ccagcctggg	ccacagtgc	840
agactccatc	tcaaaaaaaa	aagaaaagaa	aaagcctggt	taatgcacag	gtgtgagtgg	900
attgcttatg	gctatgagat	aggttgatct	cgcccttacc	ccggggtctg	gtgtatgctg	960
tgctttcttc	agcagtatgg	ctctgacatc	tcttagatgt	cccaacttca	gctgttggga	1020
gatgggtgata	ttttcaaccc	tacttcttaa	acatctgtct	ggggttccct	tagtcttgaa	1080
tgctttatgc	tcaattatgt	ggtgttgagc	ctctcttcca	caagagctcc	tccatgtttg	1140
gatagcagtt	gaagagtgtg	tgggtgggct	gttgggatga	gatggagtgt	tcagtcccca	1200
tttctcattt	tacattttaa	agtcgttctc	ccaacatagt	gtgtattggt	ctgaaggggg	1260
tgggtgggatg	ccaaagcctg	ctcaagttat	ggacattgtg	gccaccatgt	ggctttaaag	1320
attttttcta	actaataaag	tgggatatat	atttaaaaaa	aaaaaaaaaa	aaa	1373

<210> 71

<211> 1579

<212> DNA

<213> Homo sapiens

<400> 71

ggcacgagga	tttgagggg	acaaacatcc	aaaccattta	agtcacagca	ctttactccg	60
cagtgtgaat	aacacaggca	ttctcttaca	taatcacagt	acagttatca	tactctggaa	120
attgaatata	atctaataa	ctttccatac	ccagattttc	ttagatttcc	caatgatatt	180
tcttactgtc	ctcccttttag	ccttctctct	tctccattca	ggattctacc	attacatttc	240
attttcatgt	ctcttcagtc	tctcttttag	tttgttttcc	tttcttgatg	ttgccacttt	300
taggaggcca	ggccagtgtg	tttgtgaaag	atctgttctc	tttgatatgt	ttcatttttg	360
atttgtttca	ttgtttttgc	atgaatggat	tcaggctaaa	catttttggg	caggactgtt	420
tattgtatta	cctagttagt	tgttcttttc	agtcacatcat	ctggaggcac	ctgatggcag	480
ttttcccaat	attgcgaaat	taagtttgat	tatttttgta	aggtagtgtc	caccagatct	540
ctccatttta	aagacatcct	tttctcta	tactcagtgg	actgtagagt	gatgctttga	600
aactgaataa	ctaactctcc	ctaactcagt	gatttagcac	ccgttgattt	tttttttttg	660
cctgaatcaa	atattattat	agtagtttta	aatgggtgatt	ttccatttct	attcttttgt	720
tagctgccat	tcttctataa	ttttgtcttt	atattttact	gggtgttaag	attgctattc	780
cagctttcct	ttgtctttta	caccttttcc	cattctttta	tttttttcca	tccctttttg	840
tectgttttc	caatagatgg	atagaatttt	ctttctctgg	tttaaagggt	atacatttgt	900
gtgtgtgtgt	gtattctaaa	ccatttgccc	ttaaaacata	gagatgggta	ttcctgttga	960
ttaaaaaaa	ctcagtaaat	ttactatcct	ctctttaata	agatagggtac	tttatttcat	1020
tctgtgttct	tgggtgaggt	ctccctccac	ccagtcaagt	tgatgttaat	ctagaatttt	1080
tagtttttta	aattatcgat	atactttctg	tcttttctct	tttttttcac	tctctctgta	1140
tttgtgcttt	ccccctttac	tctctttccc	ttcattcctt	ctttctcact	ttcttcttct	1200
atttttcact	cttgggtaga	ttatctttta	gaaaacagca	agatattata	taattttact	1260
tatattctct	ttccaaatga	ttaaagtaat	aattaaaaat	ttttgatatg	tgtgtatgca	1320
aggataggaa	tcctcttgta	agtggaaaga	ctctaccaca	tgcattgagtc	attagtgtgt	1380
taaacactgg	gaagtggctt	taggtccagc	tgggtgctct	gaagaaggta	ggtttcttca	1440
gttctttatg	ttactgtct	ccttacctta	aaaaaggagt	gaagaatact	gactgcagag	1500
gttttgtgag	gattccggta	acacagaagc	atagaactgg	aaaagaatat	taattattgc	1560
caaaaaaaaa	aaaaaaaaaa					1579

<210> 72
 <211> 1028
 <212> DNA
 <213> Homo sapiens

<400> 72
 gcacgacaat tgaactgaac cctaaaaatg ctacttcaat tcaccttatg ggtatttggt 60
 gctatacatt tcccgaatg ccttggtatc aaagaagaat tgctaaaatg ctgtttgcaa 120
 ctgcctccta gtccaccta tgagaaggta gtatgatgac ctttggttaag ttagtacgga 180
 tttcttgaaac cacagcgccc attctaccat gtgtccaca cattgtggag ctctggattc 240
 agtgaagggg acttgaggca atttccttaa cgaaccaatt caactgtgtt atcacaaggc 300
 ttaacactta ttatccttga ctggtgagtg gttttctttt tcccgttag gtgagtggtc 360
 ggtaattctg gaatactgtc atctaaaatg gctcgtggct aaaatctacc ttcattttct 420
 gtttgaaatc taaactatat tgaagtcata aaatagaaca agaaatacag catctgttac 480
 ccagcatggt ttagctgtat tacacacaat aacagaaaag taaagcagat gcttaagttg 540
 ataaaagaag aacactcatt ataacttcta ttttaaaaag catatgaaag gttcatattc 600
 tctcatattt tcaaggcctt ttgcttttct tgttaaaaat aagatttgag aggaatttct 660
 ggtaaaactt tgggtttact catcacagc ttttcagagt aagaaaacag gcaatcgaaa 720
 aagctgtact tgtattattt acattataac aaggagcctt ttttctttc tgggaagcta 780
 tagttagtaa attgatgtaa aaaatactta gttgtattct ttacacacag ttgagaaata 840
 ttattaaat aatgcaccaa tattttataa tggattattt aaaataatgc ccatgtgctg 900
 gacacggtgg ctcatgctg taatgccagc attttggaag gccaggttg gtggatcagt 960
 tgagcctggg agtttgagac cagcctgggc aacgtggcaa aaccctgca aaaaaaaaaa 1020
 aaaaaaaaaa 1028

<210> 73
 <211> 3674
 <212> DNA
 <213> Homo sapiens

<400> 73
 ggcacgagct caaaagaat aggttgattt ttaaaggatt aataaaatc tgaatgtta 60
 agtagaagat tacattgtct agtcttgat ttctccttc tgttgctctc ttccattcac 120
 acactctcag tttctcatat ttgtagctca tttatttggt tatctcctaa gaatttgaa 180
 agtgaagcaa ctatgtgact gtattcttca ggtaaacact gactgcgctt gttggatttt 240
 ccctattttt gtgacttcaa gaataatctg cctgctgaa tacatgccat ttcacattct 300
 gaaactgggt agagtgggtt ggtgttctgc caacaattgc tagtggtgtg aattcattca 360
 tatttgccag tattgtctac ttcaaagaaa ctcttctac aagcagtgca gagctaggcc 420
 agatcaatgc tacaatcatg aagttctcat tgcattgcaat tgtgtaggat tgacaaggaa 480
 ctacagataaa aatttccagg gtgcacttcc agaaccagct tcaacatatg tctacattgc 540
 ccccaagtta ataaagtgc aaccctttac tctctcatat agccagaaat gttagaatc 600
 caaaatcttg gtgcattatt ttttcataaa cgctaaaaca tttgaagaaa caatttaatt 660
 atttaaaatt caagtatttt attcacatta ttgcaatat ccaaatgttt aaaaattccc 720
 agataattaa cttagctatta cagatctcac cttaggggtt gatgttatga agactccagt 780
 ggactgtact cacaaattga ctggacacc tatgaaagtg ggtagacctc tcagcgga 840
 ataagaaggg cttttaccta cagggcagga cagggtccca tgagagcagt tctgtggaga 900
 tataaaaaga atggaagaag gaatgcctta tagtgatatt gtgacattat atctatatat 960
 ctacatatat ctatctatct atatctacat ctatataatc ttacatttaa aattgtattc 1020
 ctacacatat tagaaactct tctaataaat gaagtaaaaa aattaaaaag aatacaaaata 1080
 ttccagcccc aaatgagaaa tcaaacatat taaaattgtt caagaaaatt tctttgaaca 1140
 cttctgaaag tttttgaaa cttagaaaaa agggaaaaaa atccagtgtt actagtaatt 1200
 tccatggtaa tacagataaa atacattctt ttaattcttg gaaattagaa aaagtggggg 1260
 gatctttcca gaaaaaacat gtgtaacatc tgcttatcac tccagctccc tctcctctc 1320
 cctctccacg ttcctttgag taaatgtctg ggaaagcatg aagcttgatg caagaaccct 1380
 gttgtactgg cgttttctc cctgtgaaa acgtaactac tgttgggagt gaattgagga 1440
 tgtagaaagg tgggtgaacc aaattgtggt caatggaaat aggagaatat ggttctcact 1500


```

cttgagaaaa aaacctaaga ttagcccagg tagttgcctg taacttcagt ttttctgcct 1560
gggtttgata tagtttaggg ttgggggttag attaagatct aaattacatc aggacaaaga 1620
gacagactat taactccaca gttaattaaag gacgtatggt ccatgtttat ttgttaaagc 1680
agtgtgaata gccttcaagc atgtgaataa tcttccatct tccccgccac acatacacac 1740
acacactttt tgtttctttc aggtagacac cttttaaaat gcagaactaa ctgaggcatt 1800
tcagtaactt tgctttcaaa tcaataaagt caaatgtatg gaaacatttt gtgccttact 1860
ctccataccc cgtgtactca aattctctac tgtatgaatt atgctttaag tagaatttcag 1920
tgccaaggag aacttggtga aataaattat tttaattttt tttttatcct ttacaaagcc 1980
atggatttta ttgtgtgat gtgtgctctg tacacaagcc atttcaatag gatggagctg 2040
ttaattattt tccaaagagt aatagacatg caaaagtttc aataaaaaact gggccattaa 2100
caagtaaat aataaactaa taagcattcc cttctaggtt tttgccaaac tgcctatcca 2160
atacaaat tgagaatcgt tgtaaaagct agttatattt cagagaaatg attttcatta 2220
ttgaaactgt tctccctagc aggccatttt ccttttttcc tgggagtta gcaagttag 2280
gagagaatag tcatgaaaag aaagggagaa agggggagaa gggagaggt taaaaagtaa 2340
gtgctcagac ctatgaacgt aatccctttg ctacaaatat ttaagagcag ctcagcttgg 2400
ttgaaactga gttttgtcat cttccatatt tgcaggaagg tattttctga cttgcaatgc 2460
agctagatgt aaaattttat tttatcatct tagaaaagcct tgactagaaa aatgaataaa 2520
tattgaggtt tctctgtcca tatctggctt gcatgtgcca gaaagcagag aatagaaaat 2580
gtaatctcca acatccaagc atcgaaaccc aaggggtagg caattctatg taggttttgg 2640
acatgaagtt tgggtgcatct tggtttatgc tggctcaact gctattaaac ctctctggct 2700
tatagtctct tcattctatt agacaagcac gtatcgaaac cttgcttcgc acaaggctct 2760
ttagttaaca atttagcagc tactgtttgt gttaaacaca cttttcacca aataggttct 2820
gaggcaaacg agagcaatga ctatttaaag aaaggctttc ccagcatcac ttacacatcc 2880
caaaactaaa aagatcaact cttccaactg agaaaagact cctggctttg aatggaaact 2940
tacagcagag agtcacagc cagggcaaca acaacgacaa caacaaacat ttggaatatt 3000
attctcaact cagcttttaa taatacatct tattattttt ctagttagaga aactacaaat 3060
cagcctcttc aacatttata tacagttaa taagcctctt gcaagtact tgttctctca 3120
cctgaggtat tttttctctc cccaccttgc cctgttctct ccttctctct tctccctttg 3180
caagagggaaa tatttaacat atttgggtcc aacttcaata atgtaataat taatacatta 3240
aaagcattta acttcctttc tagaaaaatg cacaggctaa ggcatagaca aaacaaagag 3300
aaatgctgag aaatttgcca ctggagacaa gcaatctgaa taaatatttg ccaaaagttc 3360
tttttatgtc atatatgtgc aggatttgaa ggagctatatt ttttttaatg ttgcaactag 3420
caactcatct tcggaagaca cagccaggag aatgaagtag aagtgaagg tttataaatc 3480
catttgtaag catttatccc atatatttta aattcaagaa aaattgtgtt tatctttaga 3540
atttgtatt caatacttta tgtactatgt gactcatgct tctggataaa taaagcacca 3600
aatatgtatc tgtaaccaca atcacacata ttatatataa tatatatcta tataacaaaa 3660
aaaaaaaaaa aaaa 3674

```

<210> 74

<211> 2797

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (853)

<223> n equals a,t,g, or c

<400> 74

```

ggcacgagag agcagacaga attatatgta gaggacacag gagatattta catttgtgat 60
ggagatggag gattgaataa cagattgatc aaactgtccc aagatttcat gatcctttgg 120
ctgcatggag aaaatgggac agggcctgct aagttcaaca tacctcacag tgttacactt 180
gattcagctg gtcgggtgtg ggttgcgtac cgaggaaata aaagaatcca agtatttgat 240
aaagacactg gggagtgttt aggagcatgg aataattgtt tcacagaaga ggaccttctt 300
cagtcagttt actcctgatg ggaagtactt gattgtggcc cagctgaatc ttagcaggct 360
ctcagtcgta gcagcaccac cagtgggaag cattggggag tgttctgtga tcagcacaat 420
ccaactagca gatcaagttt tgccacatct cctagaagtc gacagaaaag actggagcag 480

```

45

tctatgtagc	agaaattgga	gcaaaacaag	tacaaaaata	tgtccctttg	aatagctatg	540
ttccttcatt	tggttcataa	tgtttctttc	ccgggaatat	ttcaagtggc	agtttcagatt	600
ctcaattcac	taagtgtcta	aaaatgatgt	tcaagcacaa	gaatttattt	ttctagtata	660
aaagatctag	tatcagaaag	atttggtttt	gtatcattaa	gaatcttata	ttttgttgcc	720
ctcttgggac	ttagtttttat	ttgtaagtgc	ataaggatat	tttaatgaaa	ggaaagtaac	780
taaaaaatgg	ggttggaag	agggactaag	gtggtaacct	cattatttgc	cctggtagac	840
tgattctccc	tgngtaaaaa	aaatgggaat	aaaaatgagc	ttgcatgata	atttattaaa	900
tttcatgtga	agaaactccag	acctccagat	tgtgcaacta	acataaagt	agctgcttga	960
gagattgtaa	ataagatgaa	ctattgatta	atttgagtac	ccacagagt	ctgtgtcttg	1020
acgacttaaa	aatgaaaaag	catgattgcc	ttttgagtag	cttgagctct	agtgaggaga	1080
caagcaggca	aacagtcaca	acacagcaaa	agcgaccttg	gagcatagt	ggacttttgg	1140
agtaggagtg	ctgcatttga	ctgagggaat	catggatact	tcgcaggaga	agtgaatttg	1200
agctcagact	tgaaaactga	ggaggagctt	accaagggac	aaggaggaga	aaacaataat	1260
ttccaagtaa	agaaggtata	aaaagttaga	agtgtactgt	aaactttgat	aggcttttag	1320
gcctttttta	aagcccaact	tggtcttctg	ccattaccta	taagatattt	aatgtcagtc	1380
agccttttaa	tgtaggaata	aaatggctgg	catctaagca	ctttagtaaa	agaggttttt	1440
acaaataact	aaggatttga	gagcttctct	ctcttttttt	ttctttttct	ttcttttgg	1500
ttacatgaac	tcaacttatt	cctaactatt	gtctacctca	aagaaatttc	aagattattt	1560
agataacatg	gatattgtgc	aaatcctttg	agctgttaag	atgataattt	cctgctttcc	1620
tcctacatct	tctcctccca	ctccctcctt	tggtgtgaa	attggcttcc	caattaagac	1680
cttttttttt	tccagtttgt	tttagcttat	tataggtttt	ggaggaaact	tgccattttg	1740
taatctttca	aatcattctt	caccttctct	cacatcagct	tcctgctttt	cccagtggtt	1800
tactgtaaat	tgtgtagcat	atgacaaatc	ttgagctgac	tttcctcttc	acctgttatg	1860
gctggagtat	ttccagacc	tgaagggact	cacacttggt	ttgatacttg	gatcacatct	1920
ccgtgaggtt	aggaaggtaa	atctaccaac	aggaagccct	gtactctgta	ttccaaggcc	1980
attggtaaat	gtgttggtgc	cactgatcgg	actgtatgac	cttaaacaa	tcaccttagt	2040
tttcagtga	atgggaaat	cattgtctcc	tctttcatga	atgctgtgag	aatcagatgt	2100
gcaacaggt	catacttgcc	ctttggaaat	ctaatcacct	tggtgatacca	ttaagaggca	2160
ttttaattaa	acaaaagggc	ccttctaata	gtgctattta	tttgacaata	actatcagat	2220
ttgccttaat	tttgtgttta	tagcatttat	caaaacgtat	cctcatagac	tttatgcaga	2280
ttaatatggt	caattgattt	ggataaaaga	aagtaatttc	aggggttgg	tttaagccag	2340
gacaagaagt	gcaaatgcct	ctttgaagca	atttaggcta	aactgatttt	gaaatttcaa	2400
aatgttttat	tttactttgt	tttattaagc	caggacaaga	agtgcaaatg	cctctttgaa	2460
gcaattcagg	ctaggtaaac	cgattttgcc	atttcaaaac	gttttatttt	actttgwttt	2520
rtrtcagagt	yttawaarvc	ctgctgcaaa	tatttctgaa	tgtcttttga	aaagtgtttg	2580
ttagtgtacc	tgtgattata	gtacttctac	tttttctctt	ggattaattg	gttaaatgaa	2640
tgagaaatgt	gttatgtttt	ttactaaaaa	gtataaatta	aaatttttga	aagaaaaggc	2700
aatatttatct	ggctcccaaa	ttaaagtgtg	attttattgt	cacaaaaaaa	aaaaaaaaaa	2760
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaa			2797

<210> 75

<211> 2703

<212> DNA

<213> Homo sapiens

<400> 75

ggcacgagat	ttcctacagg	tgaaacgcc	tcattaggat	tcactgtaac	gttagtgcta	60
ttaaactcac	tagcattttt	attaatggcc	gttatctaca	ctaagctata	ctgcaacttg	120
gaaaaagagg	acctctcaga	aaactcaca	tctagcatga	ttaagcatgt	cgcttggtta	180
atcttcacca	attgcatctt	tttctgcccc	gtggcgtttt	tttcatttgc	accattgatc	240
actgcaatct	ctatcagccc	cgaataatg	aagtctgtta	ctctgatatt	ttttccattg	300
cctgcttgcc	tgaatccagt	cctgtatggt	ttcttcaacc	caaagtttaa	agaagactgg	360
aagttactga	agcgacgtgt	taccaagaaa	agtggaatcag	tttcagtttc	catcagtagc	420
caaggtggtt	gtctggaaca	ggatttctac	tacgactgtg	gcattgtactc	acatttgtag	480
ggcaacctga	ctgtttgcga	ctgctgcgaa	tcgtttcttt	taacaaagcc	agtatcatgc	540
aaacacttga	taaaatcaca	cagctgtcct	gcattggcag	tggtcttctg	ccaaagacct	600
gagggctact	ggtccgactg	tggcacacag	tcggccact	ctgattatgc	agatgaagaa	660

46

```
gattcccttg tctcagacag ttctgaccag gtgcaggcct gtggacgagc ctgcttctac 720
cagagtagag gattcccttt ggtgcgctat gcttacaatc taccaagagt taaagactga 780
actactgtgt gtgtaaccgt ttcccccgtc aaccaaaatc agtgtttata gagtgaaccc 840
tattctctac ttctatctgg gaagcacttc tgtaatcact gcctgggtgtc acttagaaga 900
aggagagggtg gcagtttatt tctcaaacca gtcattttca aagaacaggt gcctaaatta 960
taaattgggtg aaaaatgcaa tgtccaagca atgtatgac tggttgaaac aaatatatga 1020
cttgaaaaagg atcttaggtg tagtagagca atataatgtt agttttttct gatccataag 1080
aagcaaatat atacctatgt gtgtattaag cacaagataa agaacagctg ttaatatattt 1140
ttaaaaatct attttaaaat gtgattttct ataactgaag aaaatatctt gctaatttta 1200
cctaattgtt catccttaat cttaggacaa cttactgac ggccaaaaaa gggactgtcc 1260
cagctagaac tgtgagagta tacataggca ttactttatt atgttttcac ttgccatcct 1320
tgacataaga gaactataaa ttttggttaa gcaatttata aatctaaaac ctgaagatgt 1380
ttttaaaaca atattaacag ctgttaggtt aaaaaaatag ctggacattt gttttcagtc 1440
attatacatt gcttttggtc aatcagtaat tttttcttaa gtgttttggt attacactac 1500
tagaaaaaaa gtaaaaggct aattgctgtg tgggtttagt cgatttggct aaactactaa 1560
ctaattgtggg ggtttaatag tatctgaggg atttgggtgc tcatgtaat gttctcatta 1620
atgaatactt cctaatactg ttggctctac taatatcttc caatttgcgt ggatgtcacc 1680
tagcaatagc ttggattata tagaaagtaa actgtggtca atactgcat ttaattagac 1740
gaaacgggga gtaattatga cacgaagtac ttatgtttat ttcttagtga gctggattat 1800
cttgaacctg tgctattaaa tggaaatttc catacatctt cccatacta tttttataa 1860
aagagccctat tcaatagctc agagggtgaa ctctgggtta acaagataat atgttattaa 1920
taaaaataga agaagaaaga ataaagctta gtctgtgtc tttaaaaatt aaaaatttta 1980
cttgattccc atctatgggc tttagacctt ttactgggtg gagtctttaa gttataattg 2040
ttcaatatgt tttttgaaca gtgtgctaaa tcaatagcaa acccactgcc atattagtta 2100
ttctgaatat actaaaaaaa tccagctaga ttgcagttta ataattaaac tgtacatact 2160
gtgcataata tgaattttta tcttatgtaa attattttta gaacacaagt tgggaaatgt 2220
ggcttctgtt catttcgttt aattaaagct acctcctaaa ctatagtggc tgccagtagc 2280
agactgttaa attgtggttt atatactttt tgcattgtaa atagtctttg ttgtacattg 2340
tcagtgtaat aaaaacagaa tctttgtata tcaaaatcat gtagtttgta taaaatgtgg 2400
gaaggattta ttacagtgt gttgtaattt tgtaaggcca actatttaca agttttaaaa 2460
attgctatca tgtatattta cacatctgat aaatattaaa tcataacttg gtaagaaact 2520
cctaattaaa aggttttttc caaaattcag gttattgaaa atttttcatt ttattcattt 2580
aaaaactaga ataacagata tataaaagtg ttaatctttg tgctatatgg tatgaaatac 2640
aatattgtac tcagtgtttt gaattattaa agtttctaga aagcaaaaaa aaaaaaaaaa 2700
aaa
```

<210> 76
<211> 742
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (707)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (724)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (726)
<223> n equals a,t,g, or c

<400> 76

47

gcgctcgaga	atagtgggtc	ccccgggctg	caggattcgg	cacgagctca	cttcaatytct	60
tcttttgagaa	gttttttcctt	tctccgcaac	cagatgtaca	tatttgaact	ctctttgtac	120
ttggagggca	cttctttcgt	ggtagttctt	ttatttttat	taatctctgt	atccttagat	180
agtcctccaa	caaccaaagg	ttgggactct	gtcttacata	tctgggtgcc	cctcatagtg	240
cagtaataag	taagttgatt	atatacgagc	tatgtaactt	atatttttta	atggttggtat	300
atcactgagt	ttttttttt	aagaattttt	ttattgaggt	aaacttcaca	taacataaaa	360
ttaaactattt	ttaaagtgaga	agttcagtg	cacttagtat	tgtaacaat	gttgcataac	420
caccaccttt	atttaaagtt	ccaaaaaaa	tggttctctc	taaaaggaaa	ccccatccca	480
ttaagcagat	actctccatt	ccttctctcc	tccagccccc	agcaaccacc	aatctgcttt	540
ctgtctctat	ggatttatct	attcttgcta	ttttatataa	atcgaattgt	atgagacctt	600
ttgtgtctgg	cttctttcac	ttagtacaag	tttttgagat	ttatttacat	agtagcatgt	660
atcaacactt	catttttatg	gccaaataaa	attgtattat	gtgttntag	cacaaaaaaa	720
aaananaaaa	atgaccctcg	ag				742

<210> 77

<211> 1825

<212> DNA

<213> Homo sapiens

<400> 77

ggcacgagca	tgctacatgt	atacctatgt	aacacacctg	cacattctgc	acatgtatcc	60
cagaacttaa	attataataa	taaaaaaaga	ataattgggt	gatggcacat	ccagggttgc	120
caaagacagt	cccagtttat	gctgtgtgcc	tggcattatt	aataatgaca	ctgcctttaa	180
ctctcacaat	taatttggat	gataacttat	atggtaactc	tgctaaataa	aaaaaataaa	240
aattaccata	gtaacaggaa	cctacttgaa	atgatgcctc	tgtttctatt	ctggcttgaa	300
ttctgcattc	tttgaggatt	tgtagcctca	tgacagaatc	ctatctacag	gtgatgtatt	360
tcatatgatt	tttggtcatt	tttttaacaa	tctcaagccc	aataatagcc	agtgatataa	420
ggaatgtagt	tacttttctc	ccactttctg	gcaagttaag	tttagccacc	tgattacaag	480
aaggacatt	cagaggtagg	atggcacaaa	gacacagggt	ccactggaga	tcactggaag	540
cagctgcagc	agggttaaga	gaaggagtc	ccagcgagtc	ttcagtcacc	acacactaac	600
atcatcagtg	aaaagttcct	gggcctgaag	atccagctat	gttgtttcta	gttgactatt	660
ttaagtgaca	gaacttggcc	caagcattga	ccattttggt	tcctcaataa	gcctgattca	720
accagggtca	cctttgaatc	tgctctccac	ctttccaata	aacctatttt	atgcatcatt	780
cagtgcagtt	tttattttatt	tacttttttg	ctgagaaaca	tgactagatt	taggaaaaat	840
gtagaatttt	actttttttt	caatatattc	tgggttttcc	agagttttca	cgtgtttcac	900
accttctttt	gcttcccacc	attccccttt	ctatttggaa	ctagagagac	atgagtttga	960
attctagctg	tgtaacctga	gtcagttatt	taacctcttt	ttgtttctgt	ttctttgtct	1020
gtaaatagca	aaaactacaa	ttaacttttag	tcctgtctgt	acaccaaata	ttatcttgaa	1080
atattataca	tattatatgt	aattactact	gaaatgctct	aagatgccta	tgtgtgaatg	1140
gcattgttgt	aaagattaaa	taataataag	gaagtgtctg	cttcagtgtc	tggcatataa	1200
taaaagctat	tattttttacg	attattttcc	atcttataga	agaattatcg	ttcttccctt	1260
ccaaagctaa	taaatggaca	tgtgtttatc	agacagaacg	taagagctgc	caaataaata	1320
gggaataggt	gctttcgga	gtctagggaa	ataaagggtca	gggaattggt	cataaaattt	1380
agtagccata	aatagcctat	aagtagattc	cctagtttat	tctatgcagg	aaaaataaagt	1440
tctacggagc	acagattcca	aaactaattg	gtcataaata	tcacctgaaa	gtttagaaaa	1500
tgtagcatca	tggacctctt	ttcataggtt	ctaaatctta	atatctgtgg	gatggtgcag	1560
gaatctagct	ttgctaagtg	ccctcagatg	actcttgctg	ttctaggcta	aaatacatgt	1620
ggtttgctt	caatggacat	gttctggaag	aatgtttgga	tgtcacacat	tcatatttag	1680
tatgagagat	gaggtcctcc	tctcatcatt	ttcttaggtt	ctcttctctc	cactccttac	1740
cctcccatca	cttacaataa	atctttttaga	aaattagcta	tacatttggt	tcattataaa	1800
aaagaaagaa	gataaaaaaa	aaaaa				1825

<210> 78

<211> 1674

<212> DNA

<213> Homo sapiens

<400> 78

ggccacgaga	gtatctgcgg	cagctgcagg	tcctggattt	atttctcgat	tcgctgtcgg	60
aggagaatga	gaccttggtg	gagtttgcta	ttggaggcct	gtgcaacctg	tgcccagaca	120
gggccaacaa	ggagcacatc	ctgcacgcag	gaggtgtccc	actcatcatc	aactgcctat	180
ccagcccca	tgaggagacg	gtgctgtctg	ccatcaccac	gctcatgcac	ctgagcccg	240
cgggcccgcag	ctttctccca	gagctgaccg	ccacgcccg	ggtgcagtgc	atgcttcgct	300
tctccctctc	ggccagcgcc	aggttccgga	acctggcaca	gatcttcctg	gaggacttct	360
gtcccccccg	ccaagtggcc	gaggcccga	gccggcaggc	gcaattttgc	cctgggtatc	420
ccactgccga	ggagcgtggc	cccacggcag	cgctgatcca	tggagactgc	gagaccgtgg	480
cacctctact	gctggggacc	acagtcccta	tgtggacgca	gggaacgggg	agcacatact	540
gccccattgg	tgccctttca	gccatctgaa	aggcggttc	tttcagcagg	acaggcattt	600
acactgatga	aacgccactg	ggagtggaga	agccagactc	cagagacacg	gagaagatca	660
aactggagct	gcgttcatag	gctggcactc	tcaatcctac	atcaggtgcc	accaccacca	720
gactcaggcc	ctggtgtaag	aagcgcccaa	gtgcctggac	ccagaggcct	tgcaggacag	780
tgttctcagg	agctgggcct	gaggcttagg	agagctgcct	tcgctgcagg	aaatcaggga	840
ttatccctta	acagaagtgt	ctggagtagt	tttcagggtat	aggaatgaga	tgcctcgtgg	900
tgaaggatc	tcacctggg	aagatgtggt	gccccctcca	gggctctgga	ggatggatgc	960
ctcccccagg	ggctctccaa	gctgggcatt	tgggcctggt	ggatgccaac	ctggataacc	1020
tgtggcccg	cattgactgt	ccaccagcc	ttgctgttag	gcaccatgac	tccaaagatg	1080
aagatgtggt	ccctgccctt	gagtgacagc	cccagggact	taatgtggcc	atcgggcatc	1140
aagcacaagg	ccatgcaggt	gatgatacgt	cggaatagag	gcaccagccc	tggtaaactgc	1200
atcttctccc	cttgccaccc	catggccccc	gctgaaagct	tcggccctcc	tctgctgtca	1260
ctcaatgatg	gggagcccta	ccccagaagt	gtatcccacg	agggcatcag	ggagcgagt	1320
agtgttgctc	aaggagatca	ggaagagacg	gcaacgtaaa	ggatgtggct	ccatgtccat	1380
ggtgccccct	ggtcaacata	aggagcgtgg	gatccgatgg	aaaggtggag	ctcagggaaa	1440
atgggggtcc	ttgcctctcg	tgtacccct	caaggctgac	cccttagatg	gcccaggaat	1500
ggcaggtgct	acaaaaatgg	taccacgtg	ggcatggaaa	tggggcagat	taggggacca	1560
ctggactcag	aggggagggg	agggctcatc	agcaccgcct	cagggagcct	gtccctttat	1620
gttcccaaat	aaagggtcct	agaagactaa	aaaaaaaaa	aaaaaaaaa	aaaa	1674

<210> 79

<211> 2191

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1327)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1334)

<223> n equals a,t,g, or c

<400> 79

ccttctctaa	aaaagcaaac	aggcaaaact	tcatgagaat	cttgatcatg	ttaaaatttt	60
atgtccttgc	atttctccct	acacacacac	acacacacac	acacacacac	acacacacac	120
tcaacatttc	ctccacccat	atcatcactc	cttagcatct	ttattccatc	aaaactttct	180
accccttgac	attctctgtg	cagttttgaa	aattaccctc	tcagcattct	ctgttccccc	240
ccacacctag	accctgacct	ctagtcaatt	ctactaccca	ggggtgtcca	cggttccagc	300
ctcctccatg	aagcccagtt	ctatgggctc	actcctctgg	gtaagtggga	gccccagact	360
atcatcctca	ttgtatagaa	aaccaactct	gtgatgctac	ctgcccctct	tccccctctc	420
tcctgaaaga	gggctggggg	agaggtggga	ggactgggta	tggccctggc	cgggtctgta	480
ttcgtactgg	gaggagtatt	ggtactctgt	gtagaaagaa	atggggaggg	ggaaatgggg	540
tggcctcagc	atctccctaa	gtcccagcct	ttaagtcctc	ctgttcagct	tcgtcgctgc	600

49

agcttcgaga	ggagttggat	cgatcttctt	gtggaaacgt	cctcttcaat	ggttacctgc	660
cgccaacagg	taggcactcc	caatggaatg	gaggggcgcg	gaggtgggccc	aaagactaca	720
tttcccataa	ggctgcagct	ctcgggtgcc	tgtgctgtgc	gtcctgagat	acagtgaggaa	780
gtgtagttcc	ctatcagatg	cttgggctga	tgcttgaaaa	ggaagttgga	cacagcattt	840
cccatgaaac	aatgggccaa	ctaactcttg	aagctcaaaa	agatgtcctt	ggaaccccat	900
ggggaatttg	ttatcccggg	tttgggttcc	ttttgttagg	gggggctttg	ggaaaaactg	960
gggattcctc	cgatggaag	gggaaaaaat	attaaatagg	aagtatttga	cattaatgcc	1020
catgatagcc	acccactggg	gccatggaag	gtatgcccc	gtgggtattg	gaactaggct	1080
tttctgattg	gtagaagtaa	cagagtaggg	aaatttcac	tacagcttta	tttccctaac	1140
tgagtcagc	acctgtacct	tcatgaaagt	tgccagatat	aaagatctgt	agtagtactt	1200
ttccaactta	gttttatcct	gttttcccga	aaaacaatca	tttatttatt	tatttattta	1260
tttaatttta	tgagacaggg	tctggctttg	tcaccaggc	tgagtgcgag	tggtgcgac	1320
ttggctncac	tgcnacctct	gcctctcaga	ttcaagccat	ccttccacct	cagctctgcc	1380
actgagtagc	tgagactaca	agcactcgcc	accatgccc	gctaattaaa	aaaataataa	1440
tcatttttaa	tgcaagcttt	atattataaa	tacaaagtaa	acatgaaaat	aaaacccaaa	1500
catagcagtg	ttattaaact	ctggcctgta	gcagtggctc	acacctgtaa	tcctagcagt	1560
ttggaggccg	agacaggttg	attacttgag	acctggagtt	tgagaccagc	ccaggtgaca	1620
cagcaagacc	tcctctctac	taaaaataaa	aaaaaattag	ccaggtgtgg	tggtatgcac	1680
ctgtggctcc	agctacttag	gatgtgggag	tgcgaggatc	gcttgagccc	aggaggtcaa	1740
ggctgcagtg	aactatgatc	actcattaca	ccccagcctg	ggtgacagag	cgagatgctg	1800
tctcaaaaac	aaacaaaacg	aaaaacaact	ctggctagat	gctattgctt	gccaaagggtg	1860
cagtcttcca	tttattaaaa	gtgaaaatta	gggccaggca	cattggctca	tgctgttaat	1920
cccagcactt	tgaggaggctg	agggtgggtg	atcacctgag	gtcaggagtt	cgagaccagc	1980
ctggccaaca	tggtgaaacc	ttatctctgc	caaaaatata	aaagattagc	catgtgtcgt	2040
ggtgggtgct	tgtaatctca	gctacttggg	aggctgaggc	aggagaatca	cttgaaccca	2100
ggaggcagag	gttgagtgga	gccaaagattg	tgccattgca	ctccagcctg	tgcaacgagc	2160
gaaactccaa	ctcaaaaaaa	aaaaaaaaaa	a			2191

<210> 80
<211> 1335
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1287)
<223> n equals a,t,g, or c

<400> 80						
ggatatatcc	agggctgcgg	attttccccc	cttcaggttt	aatgttcct	gtttttctac	60
ctttccctcg	cagtatacgc	tcaacggcaa	gawagtggaa	gttgccgtca	aacagatcat	120
cgctggaana	gccgtggagc	aaggaggtgc	tttctcgaac	cccagagacc	tggtatctga	180
ccgggacatc	cctgagctgc	agggcttctg	agtcagactg	gctggcgtgt	cactcagccg	240
caccctgtgt	cactgttaact	tttgtgtgct	caagaaatta	tacagaaacc	tacagctgtt	300
gtaaaaggat	gctcgacca	agtgttctgt	aggcttgggg	aggatcggtt	tctctgtttt	360
gttaaatctg	gtgggtacct	ggatcttcca	cacgagtggt	attctggcct	tcagagacca	420
ggagggagtg	tctggggccg	agtgtggcac	tgtggtgaga	gtgtgtgtct	ttgcacacac	480
agtgcagcgg	gaacgggtgg	gctggctggt	gctgaagaca	gacacactcc	tgagccaagg	540
tcttgtcttc	aacctccccg	tcccgttgct	ccattttgct	ctgtgaaggt	gcaaatccct	600
ttcttccctt	cccatctcag	gctctcctgt	tttccctcag	ggccagctat	gcctttgagc	660
tttagctgtt	agaaaggaac	ccccgtgact	tgacacagct	ttcacagctg	gctgctagga	720
ccggcgggct	gggtgttcac	gtgtgtctgt	gtcatggatg	caatgcaggc	cctggaggac	780
tgctgcgtcac	ccgtcaacca	gagcgtgcct	ccgggccagc	ttccctccaa	ggaatgagtg	840
gatttcatac	aggatctctt	tattgcacag	actgaatggc	tttacctgtt	tctaattgtga	900
attagggcatg	tgaagcagtg	ggtgtccacc	cgtgtccctc	atgggtgagc	cctccagctg	960
tgagcccagg	cagtgtggtc	accgagtgag	gaccctcttc	accaggaacc	gcacccctgt	1020
gctgcctcca	cctgagagtt	gctagggggt	tcttgtcgag	atcatgtcat	cagcaccctt	1080

50

aagtcaagtc acggggtttcc atagccaggc agttgggtatg tacaattcag ttcagcgtat 1140
gaacttgtat ctctaattctg atgtccattt ttatatTTTT tgaactgag cacaatgaaa 1200
tccttttcttg aatcatttttc cttttggatt ataaaaatat gggggaaaagt gctatgatga 1260
atTTtatgca ataaatgtat acatgtntgc acatgcaccc atgctgaaaa aaaaaaaaaa 1320
aaaaaaaaaa aaaaa 1335

<210> 81
<211> 1867
<212> DNA
<213> Homo sapiens

<400> 81
cccacgcgtc cggggccacag cagagacagt ggagggcagt ggagaggacc gcgctgtcct 60
gctgtcacca agagctggag acaccatctc ccaccgagag tcatggcccc attggccctg 120
cacctectcg tcctcgtccc catcctcctc agcctgggtg cctcccagga ctggaaggct 180
gaacgcagcc aagaccctt cgagaaatgc atgcaggatc ctgactatga gcagctgctc 240
aaggtgggtga cctgggggct caatcggacc ctgaagcccc agagggtgat tgtgggtggc 300
gctgggtgtg cgggctgggt ggccgccaag gtgctcagcg atgctggaca caaggtcacc 360
atcctggagg cagataacag gatcgggggc cgcattctca cctaccggga ccagaacacg 420
ggctggattg gggagctggg agccatgcgc atgcccagct ctacagcat cctccacaag 480
ctctgccagg gcctggggct caacctgacc aagttcacc agtacgaca gaacacgtgg 540
acggaggtgc acgaagtga gctgcgcaac tatgtgggtg agaagggtgc cgagaagctg 600
ggctacgcct tgcgtcccca ggaaggggc cactcgcccc aagacatcta ccagatggct 660
ctcaaccagg ccctcaaaga cctcaaggca ctgggctgca gaaaggcgat gaagaagttt 720
gaaaggcaca cgctcttga atatcttctc ggggagggga acctgagccg gccggccgtg 780
cagcttcttg gagacgtgat gtccgaggat ggcttcttct atctcagctt cgccgaggcc 840
ctccgggccc acagctgcct cagcgacaga ctccagtaca gccgcatcgt ggggtggctg 900
gacctgctgc cgcgcgcgt gctgagctcg ctgtccgggc ttgtgctgtt gaacgcgcc 960
gtgggtggcg tgaccaggg accgcacgat gtgcacgtgc agatcgagac ctctcccccg 1020
gcgcggaatc tgaagtgct gaaggccgac gtggtgctgc tgacggcgag cggaccggcg 1080
gtgaagcgca tcaccttctc gccgcccgtg ccccgccaca tgcaggaggc gctgcccagg 1140
ctgactacg tgcgggccac caaggtgttc ctaagcttcc gcaggccctt ctggcgcgag 1200
gagcacattg aaggcgccca ctcaaacacc gatcgccgtg cgcgcattgat tttctaccg 1260
ccgcccgcgc agggcgcgct gctgctggcc tcgtacacgt ggtcggacgc ggcggcagcg 1320
ttcgccggct tgagccggga agaggcggtg cgcttggcgc tcgacgacgt ggcggcattg 1380
cacgggcctg tcgtgcgcca gctctgggac ggcaccggcg tcgtcaagcg ttggcgagg 1440
gaccagcaca gccagggtg ttttgggtga cagccgcccg cgctctggca aaccgaaaag 1500
gatgactgga cggctccctta tggccgcac tactttggcg gcgagcacac cgctaccgg 1560
cacgctggg tggagacggc ggtcaagtgt ctgcygcgcg ccatcaagat caacagccgg 1620
aaggggcctg catcggacac ggccagcccc gaggggcacg catctgacat ggaggggag 1680
gggcatgtg atggggtggc cagcagcccc tcgcatgacc tggcaaagga agaaggcagc 1740
caccctccag tccaaggcca gttatctctc caaaacacga cccacacgag gacctcgcat 1800
taaagtattt tcggaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1860
aaaaaa 1867

<210> 82
<211> 984
<212> DNA
<213> Homo sapiens

<400> 82
gaattcggca cgagcccagc ggaagccaag ccaccaggcc ccccagcgtc cagcgggagc 60
atgaacattg aggatggcgc gtgcccgcgc ctcccgtgc ccccgcgtc cgcccggtag 120
gatgtcctgg cccacgggg cattgtctt cctctggctc ttctccccc ccttggggg 180
cgggtggagg ggagtgccg tgacgtctgc cgccggaggg ggctccccgc cggccacctc 240
ctgccccgtg gcctgtcct gcagcaacca ggcagcccg gtgatctgca caggagaga 300

mctggccgag	gtcccagcca	gcatcccggg	caacacgcgg	tacctgaacc	tgcaagagaa	360
cggcatccag	gtgatccgga	cggacacgtt	caagcacctg	cggcacctgg	agattctgca	420
gctgagcaag	aacctggtgc	gcaagatcga	ggtgggcggc	ttcaacgggc	tgcccagcct	480
caacacgctg	gagctttttg	acaaccggct	gaccacgggt	cccacgcagg	ccttcgagta	540
cctgtccaag	ctgcgggagc	tctggctgcg	gaacaacccc	atcgagagca	tcccctccta	600
cgccttcaac	cgcgtgccct	cgctggcgcg	cctggacctg	ggcgagctca	agcggctgga	660
atacatctcg	gaggcgccct	tcgargggct	ggtcaacctg	cgctacctca	acctgggcat	720
gtgcaacctc	aaggacatcc	ccaactgacg	gccctgggtg	gcctggagga	gctggagctg	780
tcgggcaacc	ggctggacct	gatccgcccc	ggctccttcc	agggctctcac	cagcctgcgc	840
aagctgtggc	tcatgcacgc	ccaggtagcc	accatcgagc	gcaacgcctt	cgacgacctc	900
aagtcgctgg	aggagctcaa	cctgtccccc	aacaacctga	tgtcgctgcc	ccacgacctc	960
ttcacgcccc	tgacccgcct	cgta				984

<210> 83

<211> 2664

<212> DNA

<213> Homo sapiens

<400> 83

ggttgctggc	ccaggtgagc	gggcgcgctg	gtccaggtga	gcggggcgcgt	ccccgcgacg	60
gcgctgcctg	cccaggcgcg	ttcacgtaaa	gacagcgaga	tcctgagggc	cagccgggaa	120
ggaggcgctg	atatggagct	ggctgctgcc	aagtccgggg	cccgcgcgcg	tgccctagcgc	180
gtcctgggga	ctctgtgggg	acgcgccccg	cgcgcgcgct	cggggaccgc	tagagcccgg	240
cgtcgcgcgc	atggccctgc	tctcgcgcgc	cgcgctcacc	ctcctgctcc	tcctcatggc	300
cgtctgtgtc	aggtgccagg	agcaggccca	gaccaccgac	tggagagcca	ccctgaagac	360
catccggaac	ggcgttcata	agatagacac	gtacctgaac	gccgccttgg	acctcctggg	420
aggcgcggac	ggctctctgc	agtataaatg	catgacggat	ctaagccttt	cccacgttat	480
ggttataaac	cctccccacc	gaatggatgt	ggctctccac	tgtttggtgt	tcactctaac	540
attggtatcc	cttcctcgac	aaagtgttgc	aaccaacacg	acagggtgcta	tgaracctgt	600
ggcaaaagca	agaatgactg	tgatgaagaa	ttccagtatt	gcctctccaa	gatctgccga	660
gatgtacaga	aaacactagg	actaactcag	catgttcagg	catgtgaaac	aacagtggag	720
ctcttgtttg	acagtgttat	acatttaggt	tgtaaacat	atctggacag	ccaacgagcc	780
gcatgcagg	gtcattatga	agaaaaaact	gatctttaaa	ggagatgccg	acagctagt	840
acagatgaag	atggaagaac	ataacccttg	acaaataact	aatgttttta	caacataaaa	900
ctgtcttatt	tttgtgaaag	gattattttg	agaccttaaa	ataatttata	tcttgatgtt	960
aaaacctcaa	agcaaaaaaa	gtgagggaga	tagtgagggg	agggcacgct	tgtcttctca	1020
ggtatcttcc	ccagcattgc	tccttacttt	agtatgccaa	atgtcttgac	caatatcaaa	1080
aacaagtgtc	tgtttagcgg	agaattttga	aaagagggaat	atataactca	attttcacaa	1140
ccacatttac	caaaaaaaga	gatcaaatat	aaaattcatc	ataatgtctg	ttcaacatta	1200
tcttatttgg	aaaatgggga	aattatcact	tacaagtatt	tgtttactat	gaaattttaa	1260
atacacattt	atgcctagaa	ggaacggact	ttttttttct	attttaatta	cacataatat	1320
gtaattaaag	tacaacataa	tatgttgttt	ctctgtagcc	cgttgagcat	atgagtaagt	1380
cacatttcta	ttaggactac	ttmcaaggac	aaggtttcca	ttttccagt	tgtaaaattg	1440
gaaccatcag	ctgataacct	cgtagggagc	aaccccagga	tagctaagt	ttatgtaata	1500
tgcttagaag	gtgatgtgaa	tgcgattcag	aagcatagcc	actcccattt	tatgagctac	1560
tcacatgaca	aatgtcatct	tttgctataa	cctttgccaa	gtagagaaaa	agatggattt	1620
aatgagataa	atgaaaagat	atttamccta	atatatcaag	gcactatttg	ctgttatgct	1680
ttgttattta	tttcccagca	cttgctcctt	attgtagatt	ttttaaagac	tgtaaccttt	1740
tactaactgt	ggtcttacta	aaatttgtgc	ttgatactgc	ttttcaaaaa	gcctttaatt	1800
agagccaaaa	ggatggaaaa	ggcaagatat	aatgcctttt	tatagatctc	ttatttacat	1860
tgaaaattat	taccatatgt	ttagagcaaa	tccaagaaaa	cttcaacagc	ttctgaagat	1920
gtctatgaat	gttgaaaact	tttcaatctc	ttggaatgct	cagttatgtt	cctagaccgg	1980
tctttgctga	ctactggttg	ttaacctttc	cctagcctgg	gacctcaagc	catatatatc	2040
ctttgggtga	cccatggcca	aagttattaa	gatgaactga	ctttcaaaagt	cagagaagga	2100
cagcataggg	agaggcggtt	atttgaagt	cattacaggt	agaacagggc	agaaggaaaa	2160
gtagttctg	gagaaaaggg	catgttccca	actttggaga	tatgtcattg	ccgggaacct	2220
agtatcttcc	aacttgaatt	ggtggcagct	gttccagtga	gacaaggcac	atgtatgcct	2280

52

tgtggctaag	tgagcaaaact	gggtttccac	ttaaattgttt	gggaccctca	attgattctt	2340
tatttcaaac	ctttataaaa	ggtacagttt	tgtaagccat	tattaataat	taatgcttat	2400
cggctgggca	cagtggctca	cacctataat	cccagcactt	gggaggctga	ggcgggttga	2460
tcacttgagg	tcaggagttt	gagaccagct	ggccaacatg	gtgaaacagc	gtctctacta	2520
aaaatacaaa	aatttgccgg	gcgtgggtggc	gcatgcttat	agtctcagct	actcaggaag	2580
ctgaggtagc	agaatcactt	gaaccacagga	ggtggagggt	gcagtgcagc	gagattgtgc	2640
cactgcactg	cagcctggct	cgag				2664

<210> 84

<211> 1328

<212> DNA

<213> Homo sapiens

<400> 84

cccacgcgtc	cgggccagtg	gaggtccgca	gagtttgggc	gccaggcgag	acggcagggc	60
ttaaagttcc	gggaatcaaa	gatcaactcc	cactgaggac	aaatggacct	gtaattccgg	120
gtgtgacgag	agaacgagat	ttaccttcct	gaattaaaaa	wcwgactccc	tgcgacaagg	180
actgtgtact	gcatgaatga	ggctgagata	gttgatgttg	ctctgggaat	cctgattgag	240
agccgcaaac	aggaaaaggc	ctgcgagcag	cgggcccttg	cgggggctga	taaccagar	300
cactccccctc	cctgctccgt	gtcgccctcac	acaagtcttg	ggagcagcag	tgaggaagag	360
gacagtggga	aacaggcact	grctccaggc	ctcagccctt	cccagaggcc	gggggggttc	420
agctctgcct	gtagcaggag	ccctgaggag	gaggaggaag	aggatgtgct	gaaatacgtc	480
cgggagatct	ttttcagcta	gggcataaac	tgtgcactga	actgtctgcc	gagagcagct	540
ggaggacagc	tgagcttcca	ctggtgctgc	tgggccgccc	gcctgtggga	atggggctct	600
ctgtgctcct	acctttgtgc	cttcttgggc	ctggcagatt	cacctcaggc	cagaagcccc	660
tggaactcc	gggccttggg	gctgccgttc	tgagtgtgcg	gaaggcagga	ctcaaatga	720
gatccccattt	gactccctct	gtatgtactg	tgccctctcc	tggtctctga	ggctctggag	780
tcccaattgt	ctgtgttagt	cagtaccag	gttccaggga	aaatgatgtc	atgtggtggt	840
ccaacttact	ggaaccaag	agacagtact	ttgcaaagaa	aaggatcact	gccagggtga	900
ctggaattgc	tacagtttag	tccgcatgat	ctctcctgaa	ggaggaagcc	tggttcaaaa	960
atagtttcca	tcatgagctt	atcaatgagc	tcccacctct	ccagccagcc	tagaaagcaa	1020
acgagctgcc	cacagttctc	tgccctgtct	gggaggttga	ggccacagtg	tatagactgg	1080
taagccagac	aggcctctct	cgcgaagctg	ctaccttgct	ttcacctgta	ccttgggtccc	1140
cgggcagcta	gctataaagc	aagagggaca	ggagcccaga	agagacactg	aggacaagag	1200
atcacaccag	agtacatgtc	tctgcctctg	ttttcagtg	ggctttggac	aggaatatat	1260
gaataaatca	ctgccataca	ggttttccaa	tacacaagtg	ctagaaaata	cacacaattc	1320
cccaatga						1328

<210> 85

<211> 1342

<212> DNA

<213> Homo sapiens

<400> 85

ggccccccca	ggaggtattc	tgcctttgac	tgcaactctt	gtcgtcttat	gtgggtgttg	60
aattgatctg	tctctgcagc	cagatccagg	ctcctggaag	aaccatgtcc	ggcagctact	120
ggtcatgcca	ggcacacact	gctgcccagg	aggagctgct	gtttgaatta	tctgtgaatg	180
ttgggaagag	gaatgccaga	gctgccggct	gaaaattacc	caaccaagag	aaatctgcag	240
gatggacttt	ctggtcctct	tcttgttcta	cctggcttcg	gtgctgatgg	gtcttgttct	300
tatctgcgtc	tgctcgaaaa	cccatagctt	gaaaggcctg	gccaggggag	gagcacagat	360
attttcctgt	ataattccag	aatgtcttca	gagagccrtg	catggattgc	ttcattacct	420
tttccatacg	agaaaccaca	ccttcattgt	cctgcacctg	gtcttgcaag	ggatgggtta	480
tactgagtac	acctggggaa	gtatttggct	actgtcagga	gctggagtgt	tccttgcatt	540
accttcttct	gccctatctg	ctgctagggt	taaacctgtt	ttttttcacc	ctgacttgtg	600
gaaccaatcc	tggcattata	acaaaagcaa	atgaattatt	atttcttcat	gtttatgaat	660
ttgatgaagt	gatgtttcca	aagaacgtga	ggtgctctac	ttgtgattta	aggaaaccag	720

53

ctcgatccaa	gcactgcagt	gtgtgttaact	gggtgtgtgca	ccgttttcgac	catcactgtg	780
tttgggtgaa	caactgcatc	ggggcctgga	acatcaggta	cttcctcatc	tacgtcttga	840
ccttgacggc	ctcggctgcc	accgtcgcca	ttgtgagcac	cacttttctg	gtccacttgg	900
tggtgatgtc	agatttatac	caggagactt	acatcgatga	ccttggacac	ctccatgtta	960
tggacacggc	ctttcttatt	cagtacctgt	tcctgacttt	tcacacggatt	gtcttcatgc	1020
tgggctttgt	cgtggttctg	agcttcctcc	tgggtggcta	cctgttggtt	gtcctgtatc	1080
tggcggccac	caaccagact	actaacgagt	ggtaacagagg	tgactggggc	tggtgccagc	1140
gttgtccct	tgtggcctgg	cctccgtcag	cagagcccca	agtccaccgg	aacattcact	1200
cccattgggt	tcggagcaac	cttcaagaga	tctttctacc	tgccctttcca	tgatcatgaga	1260
ggaagaaaca	agaatgacaa	gtgtatgact	gcctttgagc	tgtagtcccc	gtttatttac	1320
acatgtggat	cctcgttttc	ca				1342

<210> 86

<211> 1154

<212> DNA

<213> Homo sapiens

<400> 86

aagacaggaa	aagctccagg	ccgtggttct	cāaagtgtgg	tccttgagca	gcagcaacat	60
cacctaggag	cctgttaggg	aaggcacagc	ctcaggccct	gccccagacc	tgacagaatca	120
gaaactctgg	ggtgaggcct	ggttatctgc	tgtaacagac	cttcacagtg	gttctgatgc	180
cctctagagc	aggagaacca	ctagcttaga	ggttgcagta	tgtttgcat	cttgccattt	240
gtgttagttc	agaggaatgg	ctgaccccca	tgtctcattt	ctaagcttca	ggcagctttt	300
ctcctgggca	gctgtcattc	tgttgagggg	aatcctgggg	actgtgggtc	ctcctccctg	360
tccgtgtgtc	cttgatctgg	cagtctaccc	ccttcatctc	ccgtggagg	ctccatgcct	420
agaggtggtc	ttcaaacaga	agaatggcaa	agataattgt	ctcgtgtttt	accctgaccc	480
cattccttta	agagggtcac	ttcttgcccc	attcatttaa	aaaccaatgt	catagttctg	540
tgattccacc	tatcagacag	tgccacgtcc	aaaggcgggg	ctctyacctc	cctgggaaga	600
gagactgttg	ctgtctgtgc	ttcctgtgtt	ctccagtccc	acgctccac	ggacccacgc	660
ccttgagagc	tcctctrgtg	tcccagggtc	tctggtgtgt	tcagagacct	ccacactcaa	720
cgaccactgg	tgctgcagaa	gggcccgtgc	ttacattcca	attaacagac	gcttttccca	780
tctaattgct	cttgcccttct	cctaacacca	cctcgggagt	gtttatgtct	attctaagtg	840
aatttctactg	tgtgaaaaaa	ttcacacctg	ttgtcccagc	gatttgggag	gccggggcgg	900
gtgtatcatt	tgagcccagg	agtttgaggc	tagcctgggc	aggatggtga	aaccccgtct	960
ctataaagaa	attttaaaaa	ttagctgggc	atagtggcac	gtgcctgtag	ttccatctac	1020
tggggaggct	ggggtgggag	gatcgcatga	gcccgggagt	ttgaggctgc	agtgagctgt	1080
gatcgagca	ctgcactcca	gtctgggcaa	cagagcaaga	ccctgtctct	taaaaaaaaa	1140
aaaaaaaaact	cgag					1154

<210> 87

<211> 1197

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (573)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1177)

<223> n equals a,t,g, or c

<220>

<221> SITE

54

<222> (1185)

<223> n equals a,t,g, or c

<400> 87

aagacaggaa	aagctccagg	ccgtggttct	caaagtgtgg	tccctggaca	gcagcaacat	60
cacctaggag	cctgttaggg	aaggcacagc	ctcaggccct	gccccagacc	tcagaaatca	120
gaaactctgg	ggtgaggcct	ggttatctgc	tgtaacagac	cttccagtgg	gttctgatgc	180
cctctagagc	aggagaacca	ctagcttaga	ggttgagta	tggttgcat	cttgccattt	240
gtgttagttc	agaggaatgg	ctgaccccca	tgctctattt	ctaagcttca	ggcagctttt	300
ctcctgggca	gctgtcattc	tggtgagggg	aatcctgggg	actgtggctc	ctcctccctg	360
tccgtgtgtc	cttgatctgg	cagtctaccc	ccttcatctc	cccgtggagg	ctccatgccc	420
agaggtgggc	ttcaaacaga	agaatggcaa	arataattgt	ctcgtgtttt	accctgacct	480
cattccttta	agagggtcac	ttcttgcccc	attcatttaa	aaaccaatgt	catagtcttg	540
tgattccacc	tatcagacag	tgccacgtcc	aangcggggc	tctcacctcc	ctgggaagag	600
agactgttgc	tgctgtgtgt	tctgtgttgc	tccagtccca	cgctcccacg	gacccacgcc	660
cttggagact	ccctcagtgt	cccagggtct	ctgggtgtgt	cagagacctc	cacactcaac	720
gacctgtggt	gctgcagaag	ggccggtgct	tacattccaa	ttacagacg	cttttcccat	780
ctaagcctc	ttgccttctc	ctaaccacc	ctcgggagtg	tttatgtcta	ttctaagtga	840
atttctactgt	gtgaaaaaat	tcacacctgt	tatcccagca	atttgggagg	ccgaggcggg	900
tgtatcattt	gggcccagga	gtttgagact	agcctgggca	agatggtgaa	accccgcttc	960
tataaagaaa	ttttaaaaaa	tggctgggca	tagtggcgcg	tgccctgtagt	tccactgtgt	1020
ggggaggctg	gggtgggagg	atcgcatgag	cccgggagtt	tgaggctgca	gtgagctgtg	1080
atcgcgccac	tgactccag	tctgggcaac	agagcaaaac	cctgtctctt	aaaaaaaaaa	1140
aaaactcgag	ggggggcccc	gtaccaatt	cgcctnats	agtgagtcg	tattaca	1197

<210> 88

<211> 910

<212> DNA

<213> Homo sapiens

<400> 88

ggcagagctg	gccttcgact	cgctatgtcc	actaacaata	tgtcggagcc	acggaggccc	60
aacaaagtgc	tgaggtgagg	accccagcgt	cgtgggcacg	ggttcgggtt	gtgggtgtgg	120
atcggggccc	tgggaagcgc	ctgtctatcc	cgggggcagg	acctgagcgc	ccctgacctt	180
cgagcctgtc	gcaggtacaa	gcccccgccg	agcgaatgta	acccggcctt	ggacgacctg	240
acgcccggact	acatgaacct	gctgggcatg	atcttcagca	tgtcgggcct	catgtcttaag	300
ctgaagtggg	gtgcttgggt	cgctgtctac	tgctccttca	tcagctttgc	caactctcgg	360
agctcggagg	acacgaagca	aatgatgagt	agcttcatgt	gagacttgcc	ctacagaaca	420
agtgaactct	gagtaagggg	tggggggacc	ccagcctggc	catcctagac	tgacacctct	480
ctcctgtctt	catgctgtcc	atctctgccg	tggtagtgct	ctatctgcag	aatcctcagc	540
ccatgacgcc	cccatggtga	taccagccta	gaagggtcac	attttggacc	ctgtctatcc	600
actaggcctg	ggctttggct	gctaaacctg	ctgccttcag	ctgccatcct	ggacttccct	660
gaatgaggcc	gtctcggtgc	ccccagctgg	atagagggaa	cctggccctt	tcctagggaa	720
caccttaggc	ttaccctctc	tgctcctctt	cccctgcctg	ctgctggggg	agatgctgtc	780
catgtttcta	ggggatttca	tttgctttct	cgttgaaacc	tggtgttaat	aaagtttttc	840
actctgaaaa	aaaaaaaaaa	aaaaaaaaac	tygrgggggg	gcccgggaacc	caattcscgg	900
gatatgagtg						910

<210> 89

<211> 1076

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1029)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1037)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1040)

<223> n equals a,t,g, or c

<400> 89

ggcacgaggg	gaaagccatg	ctcccaggac	tccttccttg	cagccttaaa	tcggctctgta	60
cggaaaattc	cgcgccttag	aaacccacgc	ttgggtgtaa	cttattattg	ttcttcctga	120
cctacttcct	gtttatcact	tccgggttca	tcattttggc	atttcggtga	tcgggttgga	180
actattgaag	cccgctttca	ggttcttttc	cccattttcc	ctttgaaagg	aagacttctg	240
gcttctccta	aatctccgtt	ctctgggtaa	ggggagtcca	agcctctgtc	atgaggaacg	300
gaaatgcgag	ggcctcgggt	gttactctaa	aatccgccct	cagcttgcac	gccggaagct	360
gcgattcctg	cagcggaaga	ggcgtgatct	ggccttcgac	tcgctatgtc	cactaacaat	420
atgtcggacc	cacggaggcc	gaacaaagtg	ctgagggtaca	agcccccgcc	gagcgaatgt	480
aacccggcct	tggacgaccc	gacgccggac	tacatgaacc	tgctgggcat	gatcttcagc	540
atgtcgggcc	tcatgcttaa	gctgaagtgg	tgtgcttggg	tcgctgtcta	ctgctccttc	600
atcagctttg	ccaactctcg	gagctcggag	gacacgaagc	aatgatgag	tagcttcatg	660
ctgtccatct	ctgccgtggt	gatgtcctat	ctgcagaatc	ctcagcccat	gacgccccca	720
tggtgatacc	agcctagaag	ggtcacatth	tggaccctgt	ctatccacta	ggcctgggct	780
ttggctgcta	aacctgctgc	cttcagctgc	catcctggac	ttccctgaat	gaggccgctc	840
cggtgccccc	agctggatag	agggaacctg	gccctttcct	agggaacacc	ctaggcttac	900
ccctcctgcc	tcccttcccc	tgcttgcctg	tgggggggat	gctgtccatg	tttctagggg	960
tattcatttg	ctttctcgth	gaaacctgth	gttaataaag	tttttcactc	tgaaaaaaaa	1020
aaaaaaaaana	raaacncgn	gggggggccc	ggaaccaat	tcsccgata	gtgagth	1076

<210> 90

<211> 1842

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (67)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (98)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (212)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1838)

<223> n equals a,t,g, or c

<400> 90

```
gcgaccgcgc ccttcagcta gctcgtcgc tcgctctgct tccctgctgc cggctgcgca 60
tggcttnggc gttggcggcg ctggcggcgg ctcgagcngc ctgcgsagcc ggtaccagca 120
gttgcagaat gaagaagagt ctggagaacc tgaacaggct gcagggtgatg ctccctccacc 180
ttacagcagc atttctgcag agagcgcaca tnattttgac tacaaggatg agtctgggtt 240
tccaaagccc ccattcttaca atgtagctac aacctgccc agttatgatg aagcggagag 300
gaccaaggct gaagctacta tccctttggt tcctgggaga gatgaggatt ttgtgggtcg 360
ggatgatttt gatgatgctg accagctgag gataggaaat gatgggattt tcatgttaac 420
ttttttcatg gcattcctct ttaactggat tgggtttttc ctgtcttttt gcctgaccac 480
ttcagctgca ggaaggatg gggccatttc aggatttggg ctctctctaa ttaaatggat 540
cctgattgtc aggtttttcca cctatttccc tggatatttt gatggtcagt actggctctg 600
gtgggtgttc cttgttttag gctttctcct gtttctcaga ggatttatca attatgcaaa 660
agttcgaag atgccagaaa ctttctcaaa tctccccagg accagagtct tctttattta 720
ttaaagatgt tttctggcaa aggccttcct gcatttatga attctctctc aagaagcaag 780
agaacacctg caggaagtga atcaagatgc agaacacaga ggaataatca cctgctttta 840
aaaaataaag tactgttgaa aagatcattt ctctctattt gttcctagggt gtaaaatttt 900
aatagttaat gcagaattct gtaatcattg aatcattagt ggttaatggt tgaaaaagct 960
cttgcaatca agtctgtgat gtattaataa tgccttatat attgtttgta gtcattttta 1020
gtagcatgag ccattgtccct gtagtcggta gggggcagtc ttgctttatt catcctccat 1080
ctcaaatga acttggaatt aaatatgtta agatatgtat aatgctggcc attttaaagg 1140
ggttttctca aaagttaaac tttgtttatg actgtgtttt tgcacataat ccatttttgc 1200
tgttcaagtt aatctagaaa tttattcaat tctgtatgaa cacctggaag caaaatcata 1260
gtgcaaaaat acatttaagg tgtgtcaaa aataagtctt taatttgtaa ataataagca 1320
ttaatttttt atagcctgta ttcacaattc tgcggtacct tattgtacct aagggattct 1380
aaaggtgttg tcaactgata aaacagaaag cactaggata caaatgaagc ttaattacta 1440
aaatgtaatt cttgacactc tttctataat tagcgttctt caccctccacc cccaccccca 1500
cccccttat tttccttttg tctcctggtg attaggccaa agtctgggag taaggagagg 1560
attaggtact taggagcaaa gaaagaagta gcttggaact tttgagatga tccctaacat 1620
actgtactac ttgcttttac aatgtgttag cagaaaccag tgggttataa tgtagaatga 1680
tgtgtttct gcccaagtgg taattcatct tggtttgcta tgtaaaaact gtaaatataa 1740
cagaacatta ataaatatct cttgtgtagc accttttaaa aaaaaaaaaa aaaaaaaaaa 1800
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaanaa aa 1842
```

<210> 91

<211> 1963

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (335)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1959)

<223> n equals a,t,g, or c

<400> 91

```
ggatcctcgc ggcggcggcg gtgcttacag cctgagaaga gcgtctcgcc cgggagcggc 60
ggcggccatc gagaccacc caaggcgcgt cccctcggc ctcccagcgc tcccaagccg 120
cagcggccgc gccctttcag ctagctcgtc cgctcgctct gcttcctcgc tcccggctgc 180
gcatggcktt ggcgttggcg gcgctggcgg cggctcgagc gcctgcgcag ccggtaccag 240
cagttgcaga atgaagaaga gtctggagaa cctgaacagg ctgcaggatga tgctcctcca 300
ccttacagca gcatttctgc agagagcgca gcatnatttt gactacaagg atgagtctgg 360
gtttccaaag ccccatctt acaatgtagc tacaacactg cccagttatg atgaagcgga 420
gaggaccaag gctgaagcta ctatcccttt gggttcctggg agagatgagg attttgtggg 480
```

57

tccgggatgat	tttcatgatg	ctgaccagct	gaggatagga	aatgatggga	ttttcatggt	540
aacttttttc	atggcattcc	tctttaactg	gattggggtt	ttcctgtctt	tttgccctgac	600
cacttcagct	gcaggaaggt	atggggccat	ttcaggattt	ggtctctctc	taattaaatg	660
gacctgatt	gtcaggtttt	ccacctat	ccctggatat	tttcatgggc	agtactggct	720
ctgggtgggtg	ttccttgttt	taggctttct	cctgtttctc	agaggattta	tcaattatgc	780
aaaagttcgg	aagatgccag	aaactttctc	aaatctcccc	aggaccagag	ttctctttat	840
ttattaaaga	tgttttctgg	caaaggcctt	cctgcattta	tgaattctct	ctcaagaagc	900
aagagaacac	ctgcaggaag	tgaatcaaga	tgcagaacac	agaggaataa	tcacctgctt	960
taaaaaaata	aagtactggt	gaaaagatca	tttctctcta	tttggtccta	ggtgtaaaat	1020
tttaatatgt	aatgcagaat	tctgtaatca	ttgaatcatt	agtgggtta	gtttgaaaaa	1080
gctcttgcaa	tcaagtctgt	gatgtattaa	taatgcctta	tatattgttt	gtagtcattt	1140
taagtagcat	gagccatgtc	cctgtagtcg	gtagggggca	gtcttgcctt	attcatcttc	1200
catctcaaaa	tgaacttggg	attaaatatt	gtaagatatg	tataatgctg	gccattttta	1260
aggggttttc	tcaaaagtta	aacttttgtt	atgactgtgt	ttttgcacat	aatccatatt	1320
tgctgttcaa	gttaattctag	aaatttatct	aattctgtat	gaacacctgg	aagcaaaatc	1380
atagtgcaca	aatacattta	aggtgtgggc	aaaaataagt	ctttaattgg	taaataataa	1440
gcattaattt	tttatagcct	gtattcacia	ttctgcggta	ccttattgta	cctaagggat	1500
tctaaagggt	ttgtcactgt	ataaaacaga	aagcactagg	atacaaatga	agcttaatta	1560
ctaaaatgta	attcttgaca	ctctttctat	aattagcgtt	cttcaccccc	acccccaccc	1620
ccacccccct	tattttcctt	ttgtctctg	gtgattaggc	caaagtctgg	gagtaaggag	1680
aggattaggt	acttaggagc	aaagaaagaa	gtagcttggg	acttttgaga	tgatccctaa	1740
catactgtac	tacttgcttt	tacaatgtgt	tagcagaaac	cagtgggtta	taattgtaga	1800
tgatgtgctt	tctgcccaag	tggttaattca	tcttggtttg	ctatgttaaa	actgtaaata	1860
caacagaaca	ttataaata	tctcttgtgt	agcaccttta	aaaaaaaaaa	aaaaaaaaaa	1920
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaa		1963

<210> 92

<211> 1487

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1470)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1487)

<223> n equals a,t,g, or c

<400> 92

gcgaccgcgc	ccctttcagc	tagctcgctc	gctcgctctg	cttccctgct	gccggctgcg	60
catggckwtg	gcgttggcgg	cgctggcggc	ggtcgagccg	gcctgcgcag	ccggtaccag	120
cagttgcaga	atgaagaaga	gtctggagaa	cctgaacagg	ctgcaggtga	tgctcctcca	180
ccttacagca	gcatttctgc	agagagcgca	gttttccacc	tatttccctg	gatattttga	240
tggtcagtac	tggtctcggg	gggtgttcct	tgtttttaggc	tttctcctgt	ttctcagagg	300
atttatcaat	tatgcaaaag	ttcgggaagat	gccagaaact	ttctcaaate	tccccaggac	360
cagagttctc	tttattttatt	aaagatgttt	tctggcaaaag	gccttccctgc	atttatgaat	420
tctctctcaa	gaagcaagag	aacacctgca	ggaagtgaat	caagatgcag	aacacagagg	480
aaataatcacc	tgcttttaaaa	aaataaagta	ctgttgaaaa	gatcatttct	ctctatttgt	540
tcctagggtg	aaaatttttaa	tagttaatgc	agaattctgt	aatcattgaa	tcattagtgg	600
ttaatgtttg	aaaaagctct	tgcaatcaag	tctgtgatgt	attaataatg	ccttatatat	660
tgttttgtagt	catttttaagt	agcatgagcc	atgtccctgt	agtcggtagg	gggcagctctt	720
gctttattca	tcctccatct	caaaatgaac	ttggaattaa	atattgtaag	atatgtataa	780
tgctggccat	tttaaaagggg	ttttctcaaa	agttaaactt	ttgttatgac	tgtgtttttg	840
cacataatcc	atatttgctg	ttcaagttaa	tctagaaatt	tattcaattc	tgtatgaaca	900

58

cctggaagca	aaatcatagt	gcaaaaatac	atttaaggtg	tggtcaaaaa	taagtcttta	960
attggtaaat	aataagcatt	aattttttat	agcctgtatt	cacaattctg	cggtacctta	1020
ttgtacctaa	gggattctaa	aggtgtgtgc	actgtataaa	acagaaagca	ctaggataca	1080
aatgaagctt	aattactaaa	atgtaattct	tgacactctt	tctataatta	gcgttcttca	1140
ccccacccc	cacccccacc	ccccttattt	tccttttgtc	tcctgggtgat	taggccaaa	1200
tctgggagta	aggagaggat	taggtactta	ggagcaaaga	aagaagtagc	ttggaacttt	1260
tgagatgatc	cctaacatac	tgactacttt	gctttttaca	tggttttagca	gaaaccagtg	1320
ggttataatg	tagaatgatg	tgctttctgc	ccaagtggta	attcatcttg	gtttgctatg	1380
ttaaaactgt	aaatacaaca	gaacattaat	aaatatctct	tggttagcac	ctttaaaaaa	1440
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	ccccgggggg	ggccccc		1487

<210> 93
 <211> 1653
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (67)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (212)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1636)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1653)
 <223> n equals a,t,g, or c

<400> 93	
gcgacgcgc	ccttcagcta
gtgcttnggc	gttggcggcg
gttgcagaat	gaagaagagt
ttacagcagc	atttctgcag
tccaaagccc	ccatcttaca
gaccaaggct	gaagctacta
ggatgatttt	gatgatgctg
ttttttcatg	gcattcctct
ttcagctgca	ggaaggatg
cctgattgtc	aggttttcca
gagaacacct	gcaggaagtg
aaaaaataaa	gtactgttga
taatagttaa	tcagaaatc
tcttgcaatc	aagtctgtga
agtagcatga	gccatgtccc
tctcaaaatg	aacttggaat
gggttttctc	aaaagttaaa
ctgttcaagt	taattctaga
agtgcaaaaa	tacatttaag
attaattttt	tatagcctgt
ggtcgcgca	gctcgtctgc
ctcgagccgc	ctcgagccgc
gcaggtgatg	ctcctccacc
agctctgggtt	tacaaggatg
aagcggagag	agttatgatg
tggtgggtcg	gatgaggatt
tcattgttaac	gatgggattt
gcctgaccac	ctgtcttttt
ttaaatggat	ctctctctaa
caagaagcaa	aattctctct
acctgcttta	aggaataatc
tgtaaaattt	tggttcctagg
ttgaaaaagc	tggttaattg
agtcatttta	tattgtttgt
tcattctcca	cttgctttat
cattttaaag	taattgctggc
ttgcacataa	gactgtgttt
tcattctcca	gacacctggg
gcaaaatcat	gctgtgtatg
aataataagc	aaataagtct
taattgtacc	ttattgtacc
taagggatcc	

59

```

taaagggtgtt gtcactgtat aaaacagaaa gcactaggat acaaataag cttaattact 1260
aaaatgtaat tcttgacact ctttctataa ttagcggttct tcacccccac ccccccaccc 1320
acccccctta ttttcctttt gtctcctggt gattaggcca aagtctggga gtaaggagag 1380
gattaggtac ttaggagcaa agaaagaagt agcttggaac ttttgagatg atccctaaca 1440
tactgtacta cttgctttta caatgtgtta gcagaaacca gtgggttata atgtagaatg 1500
atgtgctttc tgcccaagtg gtaattcatc ttggtttgct atgttaaaac tgtaaatata 1560
acagaacatt aataaatatc tcttggttag caccctttaw aaaaaaaaaa aaaaaaaaaa 1620
aaaaaaaaaa aaaaancccg ggggggggcc ccn 1653

```

```

<210> 94
<211> 1830
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (67)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (97)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (211)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1813)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1830)
<223> n equals a,t,g, or c

```

```

<400> 94
gcgaccgcgc ccttcagcta gctcgctcgc tcgctctgct tccctgctgc cggctgcgca 60
tggcttnngc gttggcggcg ctggcggcgg tcgagcngcc tgcgsagccg gtaccagcag 120
ttgcagaatg aagaagagtc tggagaacct gaacaggctg cagggtgatgc tcctccacct 180
tacagcagca tttctgcaga gagcgccatc nattttgact acaaggatga gtctgggttt 240
ccaaagcccc catcttaca tgtagctaca aactgcccc gttatgatga agcggagagg 300
accaaggctg aagctactat ccctttggtt cctgggagag atgaggattt tgtgggtcgg 360
gatgattttg atgatgctga ccagctgagg ataggaaatg atgggatttt catgttaact 420
tttttcatgg cattcctctt taactggatt ggggttttcc tgtctttttg cctgaccact 480
tcagctgcag gaaggatggt ggccatttca ggatttggtc tctctctaataaaatggatc 540
ctgattgtca ggttttccac ctatttccct ggatattttg atgggtcagta ctggctctggt 600
tgggtgttcc ttgttttagg ctttctcctg tttctcagag gatttatcaa ttatgcaaaa 660
gttcggaaga tgccagaaac tttctcaaat ctccccagga ccagagttct ctttatttat 720
taaagatgtt ttctggcaaa ggccttctg catttatgaa ttctctctca agaagcaaga 780
gaacacctgc aggaagtga tcaagatgca gaacacagag gaataatcac ctgctttaaa 840
aaaataaagt actgttgaaa agatcatttc tctctatttg ttcttaggtg taaaatttta 900
atagttaatg cagaattctg taatcattga atcattagt gttaatgttt gaaaagctc 960
ttgcaatcaa gtctgtgatg tattaataat gccttatata ttgtttgtag tcattttaag 1020

```


60

tagcatgagc	catgtccctg	tagtcggtag	ggggcagctc	tgcctttatc	atcctccatc	1080
tcaaaatgaa	cttggaaatta	aatattgtaa	gatatgtata	atgctggcca	ttttaaagg	1140
gttttctcaa	aagttaaact	tttggtatga	ctgtgttttt	gcacataatc	catatttgct	1200
gttcaagtta	atctagaaat	ttattcaatt	ctgtatgaac	acctggaagc	aaaatcatag	1260
tgcaaaaata	catttaaggt	gtggtcaaaa	ataagtcctt	aattggtaaa	taataagcat	1320
taatttttta	tagcctgtat	tcacaattct	gcggtacctt	attgtaccta	agggattcta	1380
aagggtgtgt	cactgtataa	aacagaaagc	actaggatac	aaatgaagct	taattactaa	1440
aatgtaattc	ttgacactct	ttctataatt	agcgttcttc	acccccaccc	ccacccccac	1500
cccccttatt	ttccttttgt	ctcctgggtga	ttagggccaaa	gtctgggagt	aaggagagga	1560
ttaggtactt	aggagcaaa	aaagaagtag	cttggaactt	ttgagatgat	ccctaacata	1620
ctgtactact	tgcctttaca	atgtgttagc	agaaaccagt	gggttataat	gtagaatgat	1680
gtgctttctg	ccaagtgggt	aattcatctt	ggtttgctat	gttaaaactg	taaatacaa	1740
agaaacattaa	taaatatctc	ttgtgtagca	ccttttaaaa	aaaaaaaaaa	aaaaaaaaaa	1800
aaaaaaaaaa	aancccgagg	gggggccccn				1830

<210> 95

<211> 1134

<212> DNA

<213> Homo sapiens

<400> 95

tccatctaca	gtcctcacac	aggtattcag	gaataccagg	atggcgtgcc	caagattcca	60
acagcctgta	ttacggtgga	agatgcagaa	atgatgtcaa	gaatggcttc	tcatggggtc	120
aaaattgtca	ttcagctaaa	gatgggggca	aagacctacc	cagatactga	ttccttcaac	180
actgtagcag	agatcactgg	gagcaaatat	ccagaacagg	ttgtactggt	cagtggtgac	240
ctggacagct	gggatgttgg	gcagggtgcc	atggatgatg	gcggtggagc	ctttatatca	300
tgggaagcac	tctcacttat	taaagatctt	gggctgcgtc	caaagaggac	tctgcggctg	360
gtgctctgga	ctgcagaaga	acaagggtgga	gttggtgcct	tccagtatta	tcagttacac	420
aaggtaataa	tttccaacta	cagtctgggt	atggagctctg	acgcagggaac	cttcttacc	480
actgggctgc	aattcactgg	cagtgaagag	gccaggggcat	catggaggag	gttatgagcc	540
tgctgcagcc	cctcaatata	actcaggtcc	tgagccatgg	agaagggaca	gacatcaact	600
tttggtatcca	agctggagtg	cctggagcca	gtctacttga	tgactttata	aagtatttct	660
tcttccatca	ctcccacgga	gacacatga	ctgtcatgga	tccaaagcag	atgaatgttg	720
ctgctgctgt	ttgggctggt	gtttcttatg	ttgttcgaga	catggaagaa	atgctgccta	780
ggctcctaga	acagtaagaa	agaaacgttt	tcatgcttct	ggccagggaat	cctgggtctg	840
caactttgga	aaactcctct	tcacataaca	atttcattcca	attcatcttc	aaagcacaac	900
tctatttcat	gctttctggt	attatcttct	ttgatacttt	ccaaattctc	tgcattctag	960
aaaaaggaat	cattctcccc	tcctctccac	cacatagaat	caacatatgg	tagggattac	1020
agtgggggca	tttctttata	tcacctctta	aaaacattgt	ttccacttta	aaagtaaaaca	1080
cttaataaat	ttttggaaga	tctctgaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaa	1134

<210> 96

<211> 1772

<212> DNA

<213> Homo sapiens

<400> 96

tcgacccacg	cgcccgagg	gatccccagc	cggtgcccaa	gcctgtgcct	gagcctgagc	60
ctgagcctga	gccgagccg	gagccggtcg	cgggggctcc	gggctgtggg	accgctgggc	120
ccccagcgat	ggcgaccctg	tggggaggcc	ttcttcggct	tggtctcttg	ctcagcctgt	180
cgtgcctggc	gctttccgtg	ctgctgctgg	cgactgtca	gacgcgcca	agaatttcga	240
ggatgtcaga	tgtaaatgta	tctgccctcc	ctataaagaa	aaattctggg	catattttata	300
ataagaacat	atctcagaaa	gattgtgatt	gccttcatgt	tgtggagccc	atgcctgtgc	360
gggggcctga	tgtagaagca	tactgtctac	gctgtgaatg	caaatatgaa	gaaagaagct	420
ctgtcacat	caagggtacc	attataattt	atctctccat	tttgggcctt	ctacttctgt	480
acatggtata	tcttactctg	gttgagccca	tactgaagag	gcgcctcttt	ggacatgcac	540

61

agttgatata	gagtgatgat	gatattgggg	atcaccagcc	ttttgcaaat	gcacacgatg	600
tgctagcccc	ctcccgcagt	cgagccaacg	tgctgaacaa	ggtagaatat	ggcacagcag	660
cgctggaagc	ttcaagtcca	agagcagcga	aaagtctgtc	tttgaccggc	atgttgctct	720
cagctaattg	gggaattgaa	ttcaagggtga	ctagaaagaa	acaggcagac	aactggaaaag	780
gaactgactg	ggttttgtctg	ggtttcatct	taataccttg	ttgatttcac	caactgtttgc	840
tggaagattc	aaaactggaa	gkaaaaactt	gcttgatttt	tttttcttgt	taacgtaata	900
atagagacat	ttttaaaaagc	acacagctca	aagtcagcca	ataagtcttt	tcctatttgt	960
gacttttact	aataaaaaata	aatctgcctg	taaaaataat	taaaaaatcc	tttacctgga	1020
acaagcactc	tcctttttcac	cacatagttt	taacttgact	ttccaagata	attttcaggg	1080
tttttggtgt	tggtgttttt	tggtgttttg	ttttgggtgg	agaggggagg	gatgcctggg	1140
aagtggttaa	caactttttt	caagtcactt	tactaaacaa	acttttgtaa	atagacctta	1200
ccttctattt	tcgagtttca	tttatatttt	gcagtgtagc	cagcctcatc	aaagagctga	1260
cttactcatt	tgacttttgc	actgactgta	ttatctgggt	atctgctgtg	ctgcaacttc	1320
atggtaaacg	ggatctaaaa	tgcttggtgg	cttttcacaa	aaagcagatt	ttcttcatgt	1380
actgtgatgt	ctgatgcaat	gcacatctaga	acaaactggc	catttgctag	tttactctaa	1440
agactaaaca	tagtcttggt	gtgtgtggtc	ttactcatct	tctagtacct	ttaaggacaa	1500
atcctaagga	cttggacact	tgcaataaag	aaattttatt	ttaaacccaa	gcctccctgg	1560
attgataata	tatacacatt	tgtagcatt	tccggctgtg	gtgagaggca	gctgtttgag	1620
ctccaatgtg	tgtagctttg	aactagggct	gggtgtgtg	gtgcctcttc	tgaagggtct	1680
aaccattatt	ggataactgg	ctttttttct	tcctctttgg	aatgtaacaa	taaaaataat	1740
tttgaaaca	tcaaaaaaaa	aaaaaaaaaa	aa			1772

<210> 97

<211> 2381

<212> DNA

<213> Homo sapiens

<400> 97

ccacgcgtcc	cgcaaggcca	gttctagtgt	agagagaaaa	aggagccggc	agcggctctt	60
acgcgtcccc	gggctgcgcg	ccactctctc	ggccggtaac	gcgggtgctt	gcggctgtcg	120
tcaagcgcgg	cggtgggccc	gcgggcccgg	gctgaggggc	tgccatggcg	gcggcgggccc	180
ggctccccgag	ctcctggggc	ctcttctcgc	cgctcctcgc	agggcttgca	ctactgggag	240
tcggggccggt	cccagcgcgg	gcgctgcaca	acgtcacggc	cgagctcttt	ggggcccgagg	300
cctggggcac	ccttgccggc	ttcggggacc	tcaactccga	caagcagacg	gatctcttcg	360
tgctgcccga	aagaaatgac	ttaatcgtct	ttttggcaga	ccagaatgca	ccctatttta	420
aaccctaaagt	aaaggtatct	ttcaagaatc	acagtgcatt	gataacaagt	gtagtcctcg	480
gggattatga	tgagagattct	caaatggatg	tccttctgac	atatcttccc	aaaaattatg	540
ccaagagtga	attagagagct	gttatcttct	ggggacaaaa	tcaaacatta	gatcctaaca	600
atatgaccat	actcaatagg	acttttcaag	atgagccact	aattatggat	ttcaatgggtg	660
atctaattcc	tgatattttt	ggtatcacia	atgaatccaa	ccagccacag	atactattag	720
gagggaaatt	atcatggcat	ccagcattga	ccactacaag	taaaaatgca	attccacatt	780
ctcatgcatt	tattgatctg	actgaagatt	ttacagcaga	tttattcctg	acgacattga	840
atgccaccac	tagtaccttc	cagtttgaaa	tatgggaaaa	tttggatgga	aacttctctg	900
tcagtactat	attggaaaaa	cctcaaaata	tgatgggtgt	tggaacagtc	gcatttgacg	960
actttgatgg	agatggacac	atggatcatt	tactgccagg	ctgtgaagat	aaaaattgcc	1020
aaaagagtac	catctactta	gtgagatctg	ggatgaagca	gtgggttcca	gtcctacaag	1080
atttcagcaa	taagggcaca	ctctggggct	ttgtgccatt	tgtaggatga	cagcaaccaa	1140
ctgaaatacc	aattccaatt	acccttcata	ttggagacta	caatatggat	ggctatccag	1200
acgctctggg	catactaaag	aacacatctg	gaagcaacca	gcaggccttt	ttactggaga	1260
acgtcccttg	taataatgca	agctgtgaag	aggcgctcgc	aatgtttaa	gtctactggg	1320
agctgacaga	cctaaatcaa	attaaggatg	ccatgggtgc	caccttcttt	gacatttacg	1380
aagatggaat	cttggacatt	gtagtgtctaa	gtaaaggata	tacaaagaat	gattttgcca	1440
ttcatacact	aaaaaataac	tttgaagcag	atgcttattt	tgtaaagtt	attgttctta	1500
gtgtgtctgt	ttctaattgac	tgctctcgta	gataacaccc	tttggagtga	atcaacctgg	1560
accttatatc	atgtatacaa	ctgtagatgc	aaatgggtat	ctgaaaaatg	gatcagctgg	1620
ccaactcagc	caatccgcac	atttagctct	ccaactacca	tacaacgtgc	ttggtttagg	1680
tcggagcgca	aattttcttg	accatctcta	cgttgggtatt	ccccgtccat	ctggagaaaa	1740

62

atctatacga	aaacaagagt	ggactgcaat	cattccaaat	ccccagctaa	ttgtcattcc	1800
ataccctcac	aatgtccctc	gaagttggag	tgccaaactg	tatcttacac	caagtaatat	1860
tgttctgctt	actgctatag	ctctcatcgg	tgtctgtgtt	ttcatcttgg	caataattgg	1920
cattttacat	tggcaggaaa	agaaagcaga	tgatagagaa	aaacgacaag	aagcccaccg	1980
gtttcatttt	gatgctatgt	gacttgcctt	taatattaca	taatgggaatg	gctgttcact	2040
tgattagtgt	aaacacaaat	tctggcttga	aaaaataggg	gagattaaat	attatttata	2100
aatgatgtat	cccatggtaa	ttattggaaa	gtattcaaat	aaatatgggt	tgaatatgtc	2160
acaaggtctt	ttttttttaa	gcactttgta	tataaaaatt	tgggttctct	attctgtagt	2220
gctgtacatt	ttgttcctt	tgtggaatgt	gttgcagtga	ctccagtgtt	tgtgtattta	2280
taatcttatt	tgcatcatga	tgatggaaaa	agttgtgtaa	ataaaaaata	ttaaatgagc	2340
agggaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	a		2381

<210> 98

<211> 1955

<212> DNA

<213> Homo sapiens

<400> 98

ggcacgagt	ccatccctgt	atttgtgcc	atgctcttcc	ttttctccat	ggctacactg	60
ttgaggacca	gcttcagtga	ccctggagtg	attcctcggg	cgctaccaga	tgaagcagct	120
ttcatagaaa	tggagataga	agctaccaat	ggtgcgggtg	cccaggggcca	gcgaccaccg	180
cctcgtatca	agaatttcca	gataaacac	cagattgtga	aactgaaata	ctgttacaca	240
tgcaagatct	tccggcctcc	ccgggcctcc	cattgcagca	tctgtgacaa	ctgtgtggag	300
cgcttcgacc	atcactgccc	ctgggtgggg	aattgtgttg	gaaagaggaa	ctaccgctac	360
ttctacctct	tcataccttc	tctctccctc	ctcacaatct	atgtcttcgc	cttcaacatc	420
gtctatgtgg	ccctcaaatc	tttgaataat	ggcttcttgg	agacattgaa	aggaaactcc	480
tggaactgtt	ctagaagtcc	tcatttgctt	ctttacactc	tgggtccgtcg	tgggactgac	540
tggatttcat	actttcctcg	tggctctcaa	ccagacaacc	aatgaaagac	atcaaaggat	600
catggacagg	gaagaatcgc	gtccagaatc	cctacagcca	tggcaatatt	gtgaagaact	660
gctgtgaagt	gctgtgtggc	cccttgcccc	ccagtgtgct	ggatcgaagg	ggtattttgc	720
cactggagga	aagtggaaat	cgacctccca	gtactcaaga	gaccagtagc	agcctcttgc	780
cacagagccc	agccccaca	gaacacctga	actcaaatga	gatgccggag	gacagcagca	840
ctccccgaaga	gatgccacct	ccagagcccc	cagagccacc	acaggaggca	gctgaagctg	900
agaagtagcc	tatctatgga	agagactttt	gtttgtgttt	aattagggct	atgagagatt	960
tcagggtgaga	agttaaacct	gagacagaga	gcaagtaagc	tgtccctttt	aactgttttt	1020
ctttgtctct	tagtcaccca	gttgcacact	ggcattttct	tgtctgcaagc	ttttttaaat	1080
ttctgaactc	aaggcagtgg	cagaagatgt	cagtcacctc	tgataactgg	aaaaatgggt	1140
ctcttgggcc	ctggcactgg	ttctccatgg	cctcagccac	aggggtcccct	tggaccccct	1200
ctcttccctc	cagatcccag	ccctcctgct	tggggtcact	ggctctcattc	tggggctaaa	1260
agttttcgag	actgggtcaa	atcctcccaa	gctgctgcac	gtgctgagtc	cagaggcagt	1320
cacagagacc	tctggccagg	ggatcctaac	tgggttcttg	gggtcttcag	gactgaagag	1380
gagggagagt	ggggtcagaa	gattctcctg	gccaccaagt	gccagcattg	cccacaaatc	1440
cttttaggaa	tgggacaggt	acctccact	agttgtattt	attagtgtag	cttctccttt	1500
gtctccatc	cactctgaca	ccttaagccc	cactcttttc	ccattagata	tatgtaagta	1560
gttgtagtag	agataataat	tgacatttct	cgtagactac	ccagaaactt	ttttaatacc	1620
tgtgccattc	tcaataagaa	tttatgagat	gccacgggca	tagcccttca	cactctctgt	1680
ctcatctctc	ctcctttctc	attagcccct	tttaatttgt	ttttcctttt	gactcctgct	1740
cccattagga	gcagggaatgg	cagtaataaa	agtctgcact	tgggtcattt	cttttctca	1800
gaggaagcct	gagtgtcac	ttaaacacta	tcccctcaga	ctccctgtgt	gaggcctgca	1860
gaggccctga	atgcacaaat	gggaaaccaa	ggcacagaga	ggctctcctc	tcctctcctc	1920
tcccccgatg	taccttcaaa	aaaaaaaaaa	aaaaaa			1955

<210> 99

<211> 1958

<212> DNA

<213> Homo sapiens

<400> 99

ccacgcgtcc	ggggcggttc	tggctcgtgag	aggggagccc	caggggagct	ggggcagcat	60
gactgggggtg	ataaatggcc	ggaaatttgg	cgtggccaca	ctcaacacca	gcgtgatgca	120
ggaggcacac	tccgggggtca	gcagcatcca	cagcagcatc	cgccatgtcc	cagcaaacgt	180
ggggcctctg	atgcgggtgc	tcgtggtcac	catcgccccc	atctactggg	ccctggccag	240
agagagtggg	gaagccctga	atggccactc	tctgactggg	ggcaagttcc	ggcagagtca	300
cacgtggagt	ttgctacagg	gagctgctca	cgatgaccga	gtggcccggg	gtctggatcc	360
cgatggcctc	ctgctcctcg	acgtggtggt	caatggcggt	gtccccggac	gagcctgggt	420
gacgcagatc	ttcaagtga	ggactttgaa	gaagcactac	gtgcaaaaca	gggcctggcc	480
agctgttctg	gggtccaca	cagcgcttct	tccaggggcg	cctccccctg	ttcctacgct	540
gcaaccacag	catccagtac	aacgcggccc	ggggccccc	gccccagctg	gtgcagcacc	600
tgccggcctc	agctatcagc	tcggcctttg	atccagaggc	cgaggccctg	cgcttccagc	660
tcgtacacgc	cctgcaggcg	gaggagaacg	aggtcggctg	ccccgagggc	tttgagctgg	720
actcccaggg	agcgttttgt	gtggatgtgg	acgagtgtgc	gtgggatgct	cacctctgcc	780
gagagggaca	gcgctgtgtg	aacctgtctg	ggtcctaccg	ctgcctcccc	gactgtgggc	840
ctggcttcg	gggtggctgat	ggggccggct	gtgaaaatgt	ggacgaatgc	ctggaagggg	900
ttggacgact	gtcactacaa	ccagctctgc	gagaacaccc	caggcgggtca	ccgctgcagc	960
tgccccaggg	gttaccggat	gcaggggccc	agcctgcctt	gcctagatgt	caatgagtg	1020
ctgcagctgc	ccaaggcctg	cgccctaccg	tgccacaacc	tccaggggcag	ctaccgctgc	1080
ctgtgcccc	caggccagac	cctccttcgc	gacggcaagg	cctgcacctc	actggagcgg	1140
aatggacaaa	atgtgaccac	cgtcagccac	cgaggccctc	tattgccctg	gctgcggccc	1200
tgggcctcga	tccccgggtac	ctcctaccac	gcctgggtct	ctctccgtcc	gggtcccatg	1260
gcccctgagca	gtgtggggccg	ggcctgggtg	cctcctgggt	tcacaggga	gaacggagtc	1320
tgacacagac	ttgacgagtg	ccgcgtgagg	aacctgtgtc	agcacgcctg	ccgcaacact	1380
gagggcagct	accagtgcct	gtgcccgcgc	ggctaccgtc	tgctccccag	cgggaagaac	1440
tgccaggaca	tcaacgagtg	cgaggaggag	agcatcgagt	gtggaccccg	ccagatgtgc	1500
ttcaacaccc	gtggcagcta	ccagtgtgtg	gacacaccct	gtcctgccac	ctaccggcag	1560
ggccccagcc	ctgggacgtg	cttcggggcg	tgctcgcagg	actgcggcac	ggggggccct	1620
tctacgctgc	agtaccggct	gctgcccgtg	cccctggggc	tgccgcgcca	ccacgacgtg	1680
gcccgcctca	ccgccttctc	cgaggtcggc	gtccccgcca	accgcaccga	gctcagcatg	1740
ctggagcccg	acccccgcag	ccccttcgcg	ctgcgtccgc	tgccgcggcg	ccttgccgcg	1800
gtctacaccc	gtcgcgcgct	cacccgcgcc	ggcctctacc	ggctcaccgt	gcgtgtgtgc	1860
gcaccgcgcc	accaaagcgt	cttcgtcttg	ctcatcgccg	tgtcccccta	cccctactaa	1920
acgggagagg	gcattggcgg	ccgctctaga	ggatccct			1958

<210> 100

<211> 2444

<212> DNA

<213> Homo sapiens

<400> 100

ttacgccaaag	ctggcacgag	caatgaaaga	gttaatctct	ttggctgggc	ctacagatga	60
catacagagt	acagtcccc	aggttcatgc	tttaaataac	cttagagcat	tggtcagaga	120
tacgcgcctg	ggagaaaata	ttattcctta	tggtgctgat	ggagctaagg	ctgcaattct	180
gggttttaca	tcaccgggtc	gggcagtgcg	aaattcatcc	acacttctct	ttagtgcctt	240
gatcacaaga	atttttgag	ttaaaagggc	aaaggatgaa	cattccaaaa	caaatagaat	300
gacagggaga	gagtttttct	ctcgtttccc	agaactctat	ccttttcttc	tcaaacagtt	360
ggaaactgta	gccaatagag	tagacagtga	tatgggagaa	ccaaatcgtc	atccaagcat	420
gtttctctta	cttttggtgt	tggagagact	ctacgcttcc	ccgatggatg	gtacttcttc	480
tgctctcagc	atgggacctt	ttgttccctt	cattatgagg	tgtggtcact	cacctgtcta	540
ccactcccgt	gaaatggcag	ctcgtgcctt	ggtcccatct	gttatgatag	atcacattcc	600
taataccatt	cgaactctgt	tgtccacact	ccccagctgc	actgaccagt	gtttccgggc	660
aaaaccacat	tcattggggac	acttctccag	gtttttccat	ttgttgcaag	cctactcaga	720
ctccaaaaca	cggaaacgaat	tcagacttcc	agcacgagct	gactgacatc	actgtttgta	780
ccaaagccaa	actctggctg	gccaagaggc	aaaatccatg	tttggtgacc	agagctgtat	840
atattgatata	tctcttcccta	ttgacttgct	gcctcaacag	atctgcaaa	gacaaccagg	900

64

cagttctgga	gagtccttggc	ttctgggaag	aaattcaaag	ggaattatct	caggatcaga	960
agctgataac	gggattccct	tgggccttca	aggtgccagg	cctgcccag	tacctccaga	1020
gcctcaccag	actagccatt	gctgcagtgt	gggccgccc	agccaagagt	ggagagcggg	1080
agacgaatgt	ccccatctct	ttctctcagc	tgtagaatc	tgcttccct	gaagtgcgct	1140
cactaacact	ggaagccctc	ttggaaaagt	tcttagcagc	agactctgga	cttgagaga	1200
agggcggtgc	acccttgctg	tgcaacatgg	gagagaagtt	cttattgttg	gccatgaagg	1260
aaaatcacc	agaatgcttc	tgcaagatac	tgaaaattct	acactgcatg	gacctgggtg	1320
agtggcttcc	ccagacggag	cactgtgtcc	atctgacccc	aaaggagtgc	ttgatctgga	1380
cgatggatat	tgcttccaat	gaaagatctg	aaattcagag	tgtagctctg	agacttgctt	1440
ccaaagtcac	ttcccaccac	atgcagacat	gtgtggagaa	caggggaattg	atagctgctg	1500
agctgaagca	gtgggttcag	ctgggtcatct	tgctatgtga	agaccatctt	cctacagagt	1560
ctaggctggc	cgctcggtgaa	gtcctcacca	gtactacacc	acttttcctc	accaaccccc	1620
atcctattct	tgagttgcag	gatacacttg	ctctctggaa	gtgtgtcctt	acccttctgc	1680
agagtgaagca	gcaagctgtt	agagatgcag	ccacggaaac	cgtgacaact	gccatgtcac	1740
aagaaaatac	ctgccagtc	acagagtttg	ccttctgcc	ggtggatgcc	tccatcgctc	1800
tgccctggc	cctggccgctc	ctgtgtgatc	tgctccagca	gtgggaccag	ttggcccttg	1860
gactgcccat	cctgctggga	tggtgtgttg	gagagagtga	tgacctcggtg	gcctgtgtgg	1920
agagcatgca	tcaggtggaa	gaagactacc	tgtttgaata	agcagaagtc	aacttttggtg	1980
ccgagaccct	gatctttgtg	aaatacctct	gcaagcacct	cttctgtctc	ctctcaaatg	2040
cgggtggcg	ttcccccaagc	cctgagatgc	tctgtcacct	tcaaaggatg	gtgtcagagc	2100
agtgcacact	cctgtctcag	ttcttcagag	agcttccacc	agctgctgag	tttgtgaaga	2160
cagtggagtt	cacaagacta	cgcattcaag	aggaaaggac	tttggtctgc	ttgaggtctg	2220
tgcccttttt	ggaaggaaag	gaagggaag	acaccctagt	tctcagtggt	tgggactctt	2280
atgcagaatc	gaggcagtta	actcttccaa	gaacagaagc	ggcatgttga	agaaaatctg	2340
ggggattggg	atgggggtat	gtgtggattt	ttcctccact	aaatctgcag	gaaacatggt	2400
gaacataaat	tcaaaaattt	tatcccaaaa	aaaaaaaaaa	aaaa		2444

<210> 101

<211> 2709

<212> DNA

<213> Homo sapiens

<400> 101

ggcagagat	ttcctacagg	tgaacgcc	tcattaggat	tcactgtaac	gttagtgcta	60
ttaaactcac	tagcattttt	attaatggcc	gttatctaca	ctaagctata	ctgcaacttg	120
gaaaagagg	acctctcaga	aaactcaca	tctagcatga	ttaagcatgt	cgcttggtga	180
atcttcacca	attgcattct	ttctgccc	gtggcggttt	tttcatttgc	accattgatc	240
actgcaatct	ctatcagccc	cgaataaatg	aagtctgtta	ctctgatatt	ttttccatgc	300
ctgcttgctc	gaatccagtc	ctgtatgttt	tcttcaaccc	aaagttaaaa	gaagactgga	360
agttactgaa	gacagctgtt	accaagaaaa	gtggatcagt	ttcagtttcc	atcagtagcc	420
aagtggtgtg	tctggaacag	gatttctact	acgactgtgg	catgtactca	catttgacgg	480
gcaacctgac	tgtttgccag	tgctgcgaat	cgtttctttt	aacaaagcca	gtatcatgca	540
aacacttgat	aaaatcacac	agctgtcctg	cattggcagt	ggcttcttgc	caaagacctg	600
agggctactg	gtccgactgt	ggcacacatt	cggcccactc	tgattatgca	gatgaagaag	660
attcctttgt	ctcagacagt	tctgaccagg	tgaggccctg	tgacagagcc	tgcttctacc	720
agagtagagg	attccctttg	gtgcgctatg	cttacaatct	accaagagtt	aaagactgaa	780
ctactgtgtg	tgtaaccgtt	tccccgctca	accaaaatca	gtgtttatag	agtgaacctt	840
attctcatct	ttcatctggg	aagcacttct	gtaatcactg	cctggtgtca	cttagaagaa	900
ggagaggtgg	cagtttattt	ctcaaaccag	tcattttcaa	agaacaggtg	cctaaattat	960
aaattgtgtg	aaaatgcaat	gtccaagcaa	tgtagatctt	gtttgaaaca	aatatatgac	1020
ttgaaaagga	tcttaggtgt	agtagagcaa	tataatgtta	gttttttctg	atccataaga	1080
agcaaattta	tacctatttg	tgtattaaag	acaagataaa	gaacagctgt	taatattttt	1140
taaaaattct	atttttaaaa	tgtgattttc	tataactgaa	gaaaaataac	ttgctaattt	1200
tacctaatgt	ttcatccttt	aatctcagga	caacttactg	cagggccaaa	aaagggactg	1260
tcccagctag	acctgtgaga	gtatacatag	gcattacttt	attatgtttt	cacttgccat	1320
ccttgacata	agagaactat	aaattttgtt	taagcaattt	ataaatctaa	aacctgaaga	1380
tgttttttaa	acaatattaa	cagctgttag	gttaaaaaaa	tagctggaca	tttggtttca	1440

65

gtcattatac	attgcttttg	tccaatcagt	aattttttct	taagtgtttt	gtgattacac	1500
tactagaaaa	aaagtaaaag	gctaattgct	gtgtgggttt	agtcgatttg	gctaaaactac	1560
taactaatgt	gggggtttta	tagtatctga	gggatttggg	ggcttcatgt	aatgttctca	1620
ttaatgaata	cttcctaata	tcgttggttc	tactaatatt	ttccaatttg	ctgggatgtc	1680
acctagcaat	agcttggtat	atatagaaag	taaactgtgg	tcaataacttg	catttaatta	1740
gacgaaacgg	ggagtaatta	tgacacgaag	tacttaatgt	ttatttctta	gtgagctgga	1800
ttatcttgaa	cctgtgctat	taaatggaaa	tttccataca	tcttcccat	actatttttt	1860
ataaaagagc	ctattcaata	gctcagaggt	tgaactctgg	ttaaacaaga	taatatgtta	1920
ttaataaaaa	tagaagaaga	aagaataaag	cttagtcctg	tgtcttttaa	aaattaaaaa	1980
ttttacttga	ttcccatct	atgggcttta	gacctattac	tgggtggagt	cttaaagtta	2040
taattgttca	atatgttttt	tgaacagtgt	gctaaatcaa	tagcaaaccc	actgccatat	2100
tagttattct	gaatatacta	aaaaaatcca	gctagattgc	agtttaataa	ttaaactgta	2160
catactgtgc	atataatgaa	ttttatctt	atgtaaatta	tttttagaac	acaagtggg	2220
aaatgtggct	tctgttcatt	tcgtttaatt	aaagctacct	cctaaactat	agtggctgcc	2280
agtagcagac	tgttaaatg	tgttttatat	actttttgca	ttgtaaatag	tctttgttgt	2340
acattgtcag	tgtataaaaa	acagaatctt	tgtatatcaa	aatcatgtag	tttgataaaa	2400
atgtgggaag	gatttattta	cagtgtgttg	taattttgta	aggccaacta	tttacaagtt	2460
ttaaaaattg	ctatcatgta	tatttacaca	tctgataaat	attaaatcat	aacttggtta	2520
gaaactccta	attaaaaggt	tttttccaaa	attcaggtta	ttgaaaactt	ttcattttat	2580
tcatttaaaa	actagaataa	cagatatata	aaagtgttaa	tctttgtgct	atatggtatg	2640
aaatacaata	ttgtactcag	tgttttgaat	tattaaagtt	tctagaaagc	aaaaaaaaaa	2700
aaaaaaaaaa						2709

<210> 102
 <211> 1722
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (401)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (695)
 <223> n equals a,t,g, or c

<400> 102						
gggaccgcgc	tgtcctgctg	tcaccaagag	ctggagacac	catctccac	cgagagtcac	60
ggccccattg	gccctgcacc	tcctcgtcct	cgtccccatc	ctcctcagcc	tgggtggcctc	120
ccaggactgg	aaggctgaac	gcagccaaga	ccccttcgag	aaatgcatgc	aggatcctga	180
ctatgagcag	ctgctcaagg	tcaccatcct	ggaggcagat	aacaggatcg	ggggccgcac	240
cttcacctac	cgggaccaga	wyacgggctg	gattggggag	ctgggagcca	tgcgcatgcc	300
cagctctcac	aggatcctcc	acaagctctg	ccagggcctg	gggtcaacc	tgaccaagtt	360
cacccagtac	gacaagaaca	cgtggacgga	ggtgcacgaa	ntgaagctgc	gcaactatgt	420
ggtggagaag	gtgcccagag	agctgggcta	cgccttgctg	ccccaggaaa	agggccactc	480
gcccgaagac	atctaccaga	tggctctcaa	ccaggccctc	aaagacctca	aggcaactggg	540
ctgcagaaaag	gcgatgaaga	agtttgaaag	gcacacgctc	ttggaatatc	ttctcgggga	600
ggggaacctg	agccggccgg	ccgtgcagct	tctgggagac	gtgatgtccg	aggatggctt	660
cttctatctc	agcttcgccc	aggccctccg	ggccnacagc	tgctcagcg	acagactcca	720
gtacagccgc	atcgtgggtg	gctgggacct	gctgccgcgc	gcgctgctga	gctcgtgtgc	780
cgggcttgtg	ctgttgaacg	cgcccggtgg	ggcgatgacc	cagggaccgc	acgatgtgca	840
cgtgcagatc	gagacctctc	ccccggcgcg	gaatctgaag	gtgctgaagg	ccgacgtggt	900
gctgctgacg	gcgagcggac	cggcggtgaa	gcgcacatc	ttctcgcgcg	gctgccccgc	960
cacatgcagg	aggcgctgcg	gaggctgcac	tacgtgcccg	ccaccaaggt	gttctctaagc	1020
ttccgcaggc	ccttctggcg	cgaggagcac	attgaaggcg	gccactcaaa	caccgatcgc	1080

66

```

ccgtcgcgca tgattttcta cccgccgcgc cgcgagggcg cgctgctgct ggcctcgtac 1140
acgtggtcgg acgcggcggc agcgttcgcc ggcttgagcc gggaagaggc gttgcgcttg 1200
gcgctcgacg acgtggcggc attgcacggg cctgtcgtgc gccagctctg ggacggcacc 1260
ggcgtcgtca agcgttgggc ggaggaccag cacagccagg gtggctttgt ggtacagmcg 1320
ccggcgctct ggcaaaccga aaaggatgac tggacggtcc cttatggccg catctacttt 1380
gccggcgagc acaccgccta cccgcacggc tgggtggaga cggcggtcaa gtcggcgctg 1440
cgcgccgcca tcaagatcaa cagccggaag ggcctgcat cggacacggc cagccccgag 1500
gggcacgcat ctgacatgga' ggggcagggg catgtgcatg ggggtggccag cagccccctg 1560
catgacctgg caaaggaaga aggcagccac cctccagtcc aaggccagtt atctctccaa 1620
aacacgaccc acacgaggac ctgcatttaa agtattttcg gaaaaaaaaa aaaaaaaaaa 1680
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaagggcgg cc 1722

```

<210> 103
 <211> 106
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (14)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (29)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 103
 Met Gly Ser Leu Ser Gly Cys Ala Leu Pro Phe Cys Leu Xaa Val Phe
 1 5 10 15
 Phe Leu Thr Val Ser Pro Ser Ala Val Gly Leu Leu Xaa Phe Ala Gly
 20 25 30
 Gly Pro Leu Gln Thr Leu Phe Ala Trp Val Ser Pro Val Glu Ala Ala
 35 40 45
 Glu Gln Gln Arg Leu Leu Pro Val Leu Ser Ser Gly Ser Phe Val Ser
 50 55 60
 Glu Gly Thr Cys Gln Met Pro Ala Arg Ala Leu Leu Tyr Glu Val Ser
 65 70 75 80
 Val Gly Pro Tyr Trp Glu Ile Pro Pro Ser Gln Asp Thr Arg Arg Ser
 85 90 95
 Gly Thr Tyr Leu Arg Arg Gln Ser Asp Pro
 100 105

<210> 104
 <211> 86
 <212> PRT
 <213> Homo sapiens

<400> 104
 Met Thr Leu Pro Ser Arg Ala Leu Ala Ser Leu Gly Val Gly Val Trp

67
 1 5 10 15
 Gly Met Leu Arg Leu Asn Gln Val Thr Val Ser Cys Gly Gly Ser Arg
 20 25 30
 Trp Ser Ser Arg Val Ala Leu Gly Ala Phe Ser Trp Val Cys Gly Val
 35 40 45
 Ala Leu Val Leu Gln Pro Ser Gly Gly Gly Leu Gly Leu Thr Ser Pro
 50 55 60
 Ser Glu Gly Cys Trp Glu Gly Glu Leu Ala Leu Ala Val Leu Arg Ala
 65 70 75 80
 Pro Gly Gly Ser Pro Ser
 85

 <210> 105
 <211> 302
 <212> PRT
 <213> Homo sapiens

 <400> 105
 Met Ala Arg Ala Arg Gly Ser Pro Cys Pro Pro Leu Pro Pro Gly Arg
 1 5 10 15
 Met Ser Trp Pro His Gly Ala Leu Leu Phe Leu Trp Leu Phe Ser Pro
 20 25 30
 Pro Leu Gly Ala Gly Gly Gly Gly Val Ala Val Thr Ser Ala Ala Gly
 35 40 45
 Gly Gly Ser Pro Pro Ala Thr Ser Cys Pro Val Ala Cys Ser Cys Ser
 50 55 60
 Asn Gln Ala Ser Arg Val Ile Cys Thr Arg Arg Asp Leu Ala Glu Val
 65 70 75 80
 Pro Ala Ser Ile Pro Val Asn Thr Arg Tyr Leu Asn Leu Gln Glu Asn
 85 90 95
 Gly Ile Gln Val Ile Arg Thr Asp Thr Phe Lys His Leu Arg His Leu
 100 105 110
 Glu Ile Leu Gln Leu Ser Lys Asn Leu Val Arg Lys Ile Glu Val Gly
 115 120 125
 Ala Phe Asn Gly Leu Pro Ser Leu Asn Thr Leu Glu Leu Phe Asp Asn
 130 135 140
 Arg Leu Thr Thr Val Pro Thr Gln Ala Phe Glu Tyr Leu Ser Lys Leu
 145 150 155 160
 Arg Glu Leu Trp Leu Arg Asn Asn Pro Ile Glu Ser Ile Pro Ser Tyr
 165 170 175
 Ala Phe Asn Arg Val Pro Ser Leu Arg Arg Leu Asp Leu Gly Glu Leu

68

180	185	190
Lys Arg Leu Glu Tyr Ile Ser Glu Ala Ala Phe Glu Gly Leu Val Asn		
195	200	205
Leu Arg Tyr Leu Asn Leu Gly Met Cys Asn Leu Lys Asp Ile Pro Asn		
210	215	220
Leu Thr Ala Leu Val Arg Leu Glu Glu Leu Glu Leu Ser Gly Asn Arg		
225	230	235
Leu Asp Leu Ile Arg Pro Gly Ser Phe Gln Gly Leu Thr Ser Leu Arg		
245	250	255
Lys Leu Trp Leu Met His Ala Gln Val Ala Thr Ile Glu Arg Asn Ala		
260	265	270
Phe Asp Asp Leu Lys Ser Leu Glu Glu Leu Asn Leu Ser His Asn Asn		
275	280	285
Leu Met Ser Leu Pro His Asp Leu Phe Thr Pro Leu His Arg		
290	295	300

<210> 106
 <211> 56
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (56)
 <223> Xaa equals stop translation

<400> 106
Met Pro Ser Ser Trp Leu Pro Gly Cys Phe Val Leu Leu Cys Leu Val
1 5 10 15
Ala Val Gly Cys Gln Leu Arg Glu Trp Gly Val Gly Gly Val Ser Ala
20 25 30
Val Gly Leu Leu Ala Leu Pro His Leu Gln Val Leu Gly Met Arg Gly
35 40 45
Arg Gly Leu Ile Ser Gly Gly Xaa
50 55

<210> 107
 <211> 189
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (94)
 <223> Xaa equals any of the naturally occurring L-amino acids

69

<400> 107

```

Met Ala Leu Leu Ser Arg Pro Ala Leu Thr Leu Leu Leu Leu Met
 1              5              10              15

Ala Ala Val Val Arg Cys Gln Glu Gln Ala Gln Thr Thr Asp Trp Arg
      20              25              30

Ala Thr Leu Lys Thr Ile Arg Asn Gly Val His Lys Ile Asp Thr Tyr
      35              40              45

Leu Asn Ala Ala Leu Asp Leu Leu Gly Gly Glu Asp Gly Leu Cys Gln
      50              55              60

Tyr Lys Cys Ser Asp Gly Ser Lys Pro Phe Pro Arg Tyr Gly Tyr Lys
      65              70              75              80

Pro Ser Pro Pro Asn Gly Cys Gly Ser Pro Leu Phe Gly Xaa His Leu
      85              90              95

Asn Ile Gly Ile Pro Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg
      100             105             110

Cys Tyr Glu Thr Cys Gly Lys Ser Lys Asn Asp Cys Asp Glu Glu Phe
      115             120             125

Gln Tyr Cys Leu Ser Lys Ile Cys Arg Asp Val Gln Lys Thr Leu Gly
      130             135             140

Leu Thr Gln His Val Gln Ala Cys Glu Thr Thr Val Glu Leu Leu Phe
      145             150             155             160

Asp Ser Val Ile His Leu Gly Cys Lys Pro Tyr Leu Asp Ser Gln Arg
      165             170             175

Ala Ala Cys Arg Cys His Tyr Glu Glu Lys Thr Asp Leu
      180             185

```

<210> 108

<211> 61

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (61)

<223> Xaa equals stop translation

<400> 108

```

Met Gly Asn Cys Gln Ala Gly His Asn Leu His Leu Cys Leu Ala His
 1              5              10              15

His Pro Pro Leu Val Cys Ala Thr Leu Ile Leu Leu Leu Gly Leu
      20              25              30

Ser Gly Leu Gly Leu Gly Ser Phe Leu Leu Thr His Arg Thr Gly Leu
      35              40              45

```

70

Arg Thr Leu Thr Ser Pro Arg Thr Gly Ser Leu Phe Xaa
 50 55 60

<210> 109

<211> 128

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (47)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (90)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 109

Met Arg Leu Glu Ser Leu Cys His Leu Cys Leu Ala Cys Leu Phe Phe
 1 5 10 15

Arg Leu Pro Ala Thr Arg Thr Val Tyr Cys Met Asn Glu Ala Glu Ile
 20 25 30

Val Asp Val Ala Leu Gly Ile Leu Ile Glu Ser Arg Lys Gln Xaa Lys
 35 40 45

Ala Cys Glu Gln Pro Ala Leu Ala Gly Ala Asp Asn Pro Glu His Ser
 50 55 60

Pro Pro Cys Ser Val Ser Pro His Thr Ser Ser Gly Ser Ser Ser Glu
 65 70 75 80

Glu Glu Asp Ser Gly Lys Gln Ala Leu Xaa Pro Gly Leu Ser Pro Ser
 85 90 95

Gln Arg Pro Gly Gly Ser Ser Ser Ala Cys Ser Arg Ser Pro Glu Glu
 100 105 110

Glu Glu Glu Glu Asp Val Leu Lys Tyr Val Arg Glu Ile Phe Phe Ser
 115 120 125

<210> 110

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (50)

<223> Xaa equals any of the naturally occurring L-amino acids

71

<220>

<221> SITE

<222> (69)

<223> Xaa equals stop translation

<400> 110

Met Pro His Phe Leu Asp Trp Phe Val Pro Val Tyr Leu Val Ile Ser
 1 5 10 15

Val Leu Ile Leu Val Gly Phe Gly Ala Cys Ile Tyr Tyr Phe Glu Pro
 20 25 30

Gly Leu Gln Glu Ala His Lys Trp Arg Met Gln Arg Pro Leu Val Asp
 35 40 45

Arg Xaa Leu Arg Lys Thr Leu Met Val Arg Asp Asn Leu Ala Phe Gly
 50 55 60

Gly Pro Glu Val Xaa
 65

<210> 111

<211> 123

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (123)

<223> Xaa equals stop translation

<400> 111

Met Ile Gly Gly Ile Thr Cys Ile Leu Ser Leu Ile Cys Ala Leu Ala
 1 5 10 15

Leu Ala Tyr Leu Asp Gln Arg Ala Glu Arg Ile Leu His Lys Glu Gln
 20 25 30

Gly Lys Thr Gly Glu Val Ile Lys Leu Thr Asp Val Lys Asp Phe Ser
 35 40 45

Leu Pro Leu Trp Leu Ile Phe Ile Ile Cys Val Cys Tyr Tyr Val Ala
 50 55 60

Val Phe Pro Phe Ile Gly Leu Gly Lys Val Phe Phe Thr Glu Lys Phe
 65 70 75 80

Gly Phe Ser Ser Gln Ala Ala Ser Ala Ile Asn Ser Val Val Tyr Val
 85 90 95

Ile Ser Ala Pro Met Ser Pro Val Phe Gly Leu Leu Val Asp Lys Thr
 100 105 110

Gly Lys Asn Ile Ile Trp Val Leu Cys Ala Xaa
 115 120

72

<210> 112
 <211> 83
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (83)
 <223> Xaa equals stop translation

<400> 112
 Met Glu Lys Gln Cys Cys Ser His Pro Val Ile Cys Ser Leu Ser Thr
 1 5 10 15
 Met Tyr Thr Phe Leu Leu Gly Ala Ile Phe Ile Ala Leu Ser Ser Ser
 20 25 30
 Arg Ile Leu Leu Val Lys Tyr Ser Ala Asn Glu Gly Lys Leu Arg Leu
 35 40 45
 Gly Ile Cys Met Glu His Phe His Leu Ile Thr His Leu Ser Leu Ala
 50 55 60
 Phe Gly Ser Val Ile Tyr Asn Met Glu Ile Ile Met Pro Phe Ala Ser
 65 70 75 80
 Cys Glu Xaa

<210> 113
 <211> 345
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (53)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (345)
 <223> Xaa equals stop translation

<400> 113
 Met Asp Phe Leu Val Leu Phe Leu Phe Tyr Leu Ala Ser Val Leu Met
 1 5 10 15
 Gly Leu Val Leu Ile Cys Val Cys Ser Lys Thr His Ser Leu Lys Gly
 20 25 30
 Leu Ala Arg Gly Gly Ala Gln Ile Phe Ser Cys Ile Ile Pro Glu Cys
 35 40 45
 Leu Gln Arg Ala Xaa His Gly Leu Leu His Tyr Leu Phe His Thr Arg
 50 55 60

73

Asn His Thr Phe Ile Val Leu His Leu Val Leu Gln Gly Met Val Tyr
 65 70 75 80
 Thr Glu Tyr Thr Trp Glu Val Phe Gly Tyr Cys Gln Glu Leu Glu Leu
 85 90 95
 Ser Leu His Tyr Leu Leu Leu Pro Tyr Leu Leu Leu Gly Val Asn Leu
 100 105 110
 Phe Phe Phe Thr Leu Thr Cys Gly Thr Asn Pro Gly Ile Ile Thr Lys
 115 120 125
 Ala Asn Glu Leu Leu Phe Leu His Val Tyr Glu Phe Asp Glu Val Met
 130 135 140
 Phe Pro Lys Asn Val Arg Cys Ser Thr Cys Asp Leu Arg Lys Pro Ala
 145 150 155 160
 Arg Ser Lys His Cys Ser Val Cys Asn Trp Cys Val His Arg Phe Asp
 165 170 175
 His His Cys Val Trp Val Asn Asn Cys Ile Gly Ala Trp Asn Ile Arg
 180 185 190
 Tyr Phe Leu Ile Tyr Val Leu Thr Leu Thr Ala Ser Ala Ala Thr Val
 195 200 205
 Ala Ile Val Ser Thr Thr Phe Leu Val His Leu Val Val Met Ser Asp
 210 215 220
 Leu Tyr Gln Glu Thr Tyr Ile Asp Asp Leu Gly His Leu His Val Met
 225 230 235 240
 Asp Thr Val Phe Leu Ile Gln Tyr Leu Phe Leu Thr Phe Pro Arg Ile
 245 250 255
 Val Phe Met Leu Gly Phe Val Val Val Leu Ser Phe Leu Leu Gly Gly
 260 265 270
 Tyr Leu Leu Phe Val Leu Tyr Leu Ala Ala Thr Asn Gln Thr Thr Asn
 275 280 285
 Glu Trp Tyr Arg Gly Asp Trp Ala Trp Cys Gln Arg Cys Pro Leu Val
 290 295 300
 Ala Trp Pro Pro Ser Ala Glu Pro Gln Val His Arg Asn Ile His Ser
 305 310 315 320
 His Gly Leu Arg Ser Asn Leu Gln Glu Ile Phe Leu Pro Ala Phe Pro
 325 330 335
 Cys His Glu Arg Lys Lys Gln Glu Xaa
 340 345

<210> 114

<211> 181

<212> PRT

74

<213> Homo sapiens

<220>

<221> SITE

<222> (110)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 114

Met Ala Asp Pro His Val Ser Phe Leu Ser Phe Arg Gln Leu Phe Ser
 1 5 10 15

Trp Ala Ala Val Ile Leu Leu Arg Gly Ile Leu Gly Thr Val Ala Pro
 20 25 30

Pro Pro Cys Pro Cys Val Leu Asp Leu Ala Val Tyr Pro Leu His Leu
 35 40 45

Pro Val Glu Ala Pro Cys Leu Glu Val Val Phe Lys Gln Lys Asn Gly
 50 55 60

Lys Asp Asn Cys Leu Val Phe Tyr Pro Asp Pro Ile Pro Leu Arg Gly
 65 70 75 80

Ser Leu Leu Gly Pro Phe Ile Lys Asn Gln Cys His Ser Ser Val Ile
 85 90 95

Pro Leu Ser Asp Ser Ala Thr Ser Lys Ala Arg Ala Leu Xaa Leu Pro
 100 105 110

Gly Arg Glu Thr Val Leu Ser Val Leu Pro Val Phe Ser Ser Pro Thr
 115 120 125

Leu Pro Arg Thr His Ala Leu Gly Asp Ser Leu Gly Val Pro Gly Leu
 130 135 140

Leu Val Cys Ser Glu Thr Ser Thr Leu Asn Asp His Trp Cys Cys Arg
 145 150 155 160

Arg Ala Gly Ala Tyr Ile Pro Ile Asn Arg Arg Phe Ser His Leu Met
 165 170 175

Pro Leu Ala Phe Ser
 180

<210> 115

<211> 116

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (116)

<223> Xaa equals stop translation

<400> 115

Met Pro Ser Ser Ser Ser Gly Leu Gly Ser Pro Ser Arg Pro Pro Ser
 1 5 10 15

75

Ser Phe Leu Cys Leu Leu Leu Leu Leu Pro Pro Ala Ala Leu Ala
 20 25 30
 Leu Leu Leu Phe Phe Leu Asp Phe Phe Pro Pro Arg Ala Ala Val Ser
 35 40 45
 Pro Phe Leu Pro Asp His Cys Ser Ala Arg Gln Pro Arg Val Trp Arg
 50 55 60
 Arg Glu Thr Leu Asn Arg Ser Ala Ser Gly Leu Gly Cys Trp Ala Arg
 65 70 75 80
 Ser Thr Glu Gln Gly Ala Val Gly Val Ala Thr Gly Thr Val Leu Asp
 85 90 95
 Ile Ser Leu Pro Ala Ser Cys Leu Ser Leu Trp Pro Pro Gly Pro Ser
 100 105 110
 Gly Gly Ile Xaa
 115

<210> 116
 <211> 71
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (71)
 <223> Xaa equals stop translation

<400> 116
 Met Asn Leu Leu Gly Met Ile Phe Ser Met Cys Gly Leu Met Leu Lys
 1 5 10 15
 Leu Lys Trp Cys Ala Trp Val Ala Val Tyr Cys Ser Phe Ile Ser Phe
 20 25 30
 Ala Asn Ser Arg Ser Ser Glu Asp Thr Lys Gln Met Met Ser Ser Phe
 35 40 45
 Met Leu Ser Ile Ser Ala Val Val Met Ser Tyr Leu Gln Asn Pro Gln
 50 55 60
 Pro Met Thr Pro Pro Trp Xaa
 65 70

<210> 117
 <211> 64
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (64)

76

<223> Xaa equals stop translation

<400> 117

Met Arg Asp Leu Ser Phe Leu Tyr Thr Leu Leu Trp Leu Pro Glu Ile
 1 5 10 15

Trp Gln Ala Leu Ala Gly Gly Ile Arg Leu Asp Glu Val Glu Leu Leu
 20 25 30

Glu Asn Glu Ala Val Leu Gly Glu Glu Met Arg Leu Tyr Arg Lys Ile
 35 40 45

Asn Glu Val Val Leu Ser Gly Asn Glu Val Val Leu Gly Gly Lys Xaa
 50 55 60

<210> 118

<211> 335

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (335)

<223> Xaa equals stop translation

<400> 118

Met Gly Ile Phe Pro Gly Ile Ile Leu Ile Phe Leu Arg Val Lys Phe
 1 5 10 15

Ala Thr Ala Ala Val Ile Val Ser Gly Val Ser Lys His Leu His Cys
 20 25 30

Ile Ser His Gln Lys Ser Thr Thr Val Ser His Glu Met Ser Gly Leu
 35 40 45

Asn Trp Lys Pro Phe Val Tyr Gly Gly Leu Ala Ser Ile Val Ala Glu
 50 55 60

Phe Gly Thr Phe Pro Val Asp Leu Thr Lys Thr Arg Leu Gln Val Gln
 65 70 75 80

Gly Gln Ser Ile Asp Ala Arg Phe Lys Glu Ile Lys Tyr Arg Gly Met
 85 90 95

Phe His Ala Leu Phe Arg Ile Cys Lys Glu Glu Gly Val Leu Ala Leu
 100 105 110

Tyr Ser Gly Ile Ala Pro Ala Leu Leu Arg Gln Ala Ser Tyr Gly Thr
 115 120 125

Ile Lys Ile Gly Ile Tyr Gln Ser Leu Lys Arg Leu Phe Val Glu Arg
 130 135 140

Leu Glu Asp Glu Thr Leu Leu Ile Asn Met Ile Cys Gly Val Val Ser

77

145 150 155 160
 Gly Val Ile Ser Ser Thr Ile Ala Asn Pro Thr Asp Val Leu Lys Ile
 165 170 175
 Arg Met Gln Ala Gln Gly Ser Leu Phe Gln Gly Ser Met Ile Gly Ser
 180 185 190
 Phe Ile Asp Ile Tyr Gln Gln Glu Gly Thr Arg Gly Leu Trp Arg Gly
 195 200 205
 Val Val Pro Thr Ala Gln Arg Ala Ala Ile Val Val Gly Val Glu Leu
 210 215 220
 Pro Val Tyr Asp Ile Thr Lys Lys His Leu Ile Leu Ser Gly Met Met
 225 230 235 240
 Gly Asp Thr Ile Leu Thr His Phe Val Ser Ser Phe Thr Cys Gly Leu
 245 250 255
 Ala Gly Ala Leu Ala Ser Asn Pro Val Asp Val Val Arg Thr Arg Met
 260 265 270
 Met Asn Gln Arg Ala Ile Val Gly His Val Asp Leu Tyr Lys Gly Thr
 275 280 285
 Val Asp Gly Ile Leu Lys Met Trp Lys His Glu Gly Phe Phe Ala Leu
 290 295 300
 Tyr Lys Gly Phe Trp Pro Asn Trp Leu Arg Leu Gly Pro Trp Asn Ile
 305 310 315 320
 Ile Phe Phe Ile Thr Tyr Glu Gln Leu Lys Arg Leu Gln Ile Xaa
 325 330 335

<210> 119

<211> 221

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (51)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 119

Met Ala Leu Ala Leu Ala Ala Leu Ala Ala Val Glu Pro Ala Cys Gly
 1 5 10 15

Ser Arg Tyr Gln Gln Leu Gln Asn Glu Glu Glu Ser Gly Glu Pro Glu
 20 25 30

Gln Ala Ala Gly Asp Ala Pro Pro Pro Tyr Ser Ser Ile Ser Ala Glu
 35 40 45

Ser Ala Xaa Tyr Phe Asp Tyr Lys Asp Glu Ser Gly Phe Pro Lys Pro
 50 55 60

78

Pro Ser Tyr Asn Val Ala Thr Thr Leu Pro Ser Tyr Asp Glu Ala Glu
 65 70 75 80
 Arg Thr Lys Ala Glu Ala Thr Ile Pro Leu Val Pro Gly Arg Asp Glu
 85 90 95
 Asp Phe Val Gly Arg Asp Asp Phe Asp Asp Ala Asp Gln Leu Arg Ile
 100 105 110
 Gly Asn Asp Gly Ile Phe Met Leu Thr Phe Phe Met Ala Phe Leu Phe
 115 120 125
 Asn Trp Ile Gly Phe Phe Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala
 130 135 140
 Gly Arg Tyr Gly Ala Ile Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp
 145 150 155 160
 Ile Leu Ile Val Arg Phe Ser Thr Tyr Phe Pro Gly Tyr Phe Asp Gly
 165 170 175
 Gln Tyr Trp Leu Trp Trp Val Phe Leu Val Leu Gly Phe Leu Leu Phe
 180 185 190
 Leu Arg Gly Phe Ile Asn Tyr Ala Lys Val Arg Lys Met Pro Glu Thr
 195 200 205
 Phe Ser Asn Leu Pro Arg Thr Arg Val Leu Phe Ile Tyr
 210 215 220

<210> 120
 <211> 473
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (473)
 <223> Xaa equals stop translation

<400> 120
 Met Lys Phe Leu Ile Phe Ala Phe Phe Gly Gly Val His Leu Leu Ser
 1 5 10 15
 Leu Cys Ser Gly Lys Ala Ile Cys Lys Asn Gly Ile Ser Lys Arg Thr
 20 25 30
 Phe Glu Glu Ile Lys Glu Glu Ile Ala Ser Cys Gly Asp Val Ala Lys
 35 40 45
 Ala Ile Ile Asn Leu Ala Val Tyr Gly Lys Ala Gln Asn Arg Ser Tyr
 50 55 60
 Glu Arg Leu Ala Leu Leu Val Asp Thr Val Gly Pro Arg Leu Ser Gly
 65 70 75 80

79

Ser Lys Asn Leu Glu Lys Ala Ile Gln Ile Met Tyr Gln Asn Leu Gln
 85 90 95
 Gln Asp Gly Leu Glu Lys Val His Leu Glu Pro Val Arg Ile Pro His
 100 105 110
 Trp Glu Arg Gly Glu Glu Ser Ala Val Met Leu Glu Pro Arg Ile His
 115 120 125
 Lys Ile Ala Ile Leu Gly Leu Gly Ser Ser Ile Gly Thr Pro Pro Glu
 130 135 140
 Gly Ile Thr Ala Glu Val Leu Val Val Thr Ser Phe Asp Glu Leu Gln
 145 150 155 160
 Arg Arg Ala Ser Glu Ala Arg Gly Lys Ile Val Val Tyr Asn Gln Pro
 165 170 175
 Tyr Ile Asn Tyr Ser Arg Thr Val Gln Tyr Arg Thr Gln Gly Ala Val
 180 185 190
 Glu Ala Ala Lys Val Gly Ala Leu Ala Ser Leu Ile Arg Ser Val Ala
 195 200 205
 Ser Phe Ser Ile Tyr Ser Pro His Thr Gly Ile Gln Glu Tyr Gln Asp
 210 215 220
 Gly Val Pro Lys Ile Pro Thr Ala Cys Ile Thr Val Glu Asp Ala Glu
 225 230 235 240
 Met Met Ser Arg Met Ala Ser His Gly Ile Lys Ile Val Ile Gln Leu
 245 250 255
 Lys Met Gly Ala Lys Thr Tyr Pro Asp Thr Asp Ser Phe Asn Thr Val
 260 265 270
 Ala Glu Ile Thr Gly Ser Lys Tyr Pro Glu Gln Val Val Leu Val Ser
 275 280 285
 Gly His Leu Asp Ser Trp Asp Val Gly Gln Gly Ala Met Asp Asp Gly
 290 295 300
 Gly Gly Ala Phe Ile Ser Trp Glu Ala Leu Ser Leu Ile Lys Asp Leu
 305 310 315 320
 Gly Leu Arg Pro Lys Arg Thr Leu Arg Leu Val Leu Trp Thr Ala Glu
 325 330 335
 Glu Gln Gly Gly Val Gly Ala Phe Gln Tyr Tyr Gln Leu His Lys Val
 340 345 350
 Asn Ile Ser Asn Tyr Ser Leu Val Met Glu Ser Asp Ala Gly Thr Phe
 355 360 365
 Leu Pro Thr Gly Leu Gln Phe Thr Gly Ser Glu Lys Ala Arg Ala Ile
 370 375 380
 Met Glu Glu Val Met Ser Leu Leu Gln Pro Leu Asn Ile Thr Gln Val

80

385 390 395 400
 Leu Ser His Gly Glu Gly Thr Asp Ile Asn Phe Trp Ile Gln Ala Gly
 405 410 415
 Val Pro Gly Ala Ser Leu Leu Asp Asp Leu Tyr Lys Tyr Phe Phe Phe
 420 425 430
 His His Ser His Gly Asp Thr Met Thr Val Met Asp Pro Lys Gln Met
 435 440 445
 Asn Val Ala Ala Ala Val Trp Ala Val Val Ser Tyr Val Val Ala Asp
 450 455 460
 Met Glu Glu Met Leu Pro Arg Ser Xaa
 465 470

<210> 121
 <211> 168
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (168)
 <223> Xaa equals stop translation

<400> 121
 Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser
 1 5 10 15
 Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr
 20 25 30
 Pro Pro Ser Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly
 35 40 45
 Pro Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu
 50 55 60
 Arg Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile
 65 70 75 80
 Leu Gly Leu Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro
 85 90 95
 Ile Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp
 100 105 110
 Asp Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu
 115 120 125
 Ala Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Ala
 130 135 140
 Gln Gln Arg Trp Lys Leu Gln Val Gln Glu Gln Arg Lys Ser Val Phe
 145 150 155 160

81

Asp Arg His Val Val Leu Ser Xaa
165

<210> 122
<211> 47
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (47)
<223> Xaa equals stop translation

<400> 122
Met Lys Phe Ile Leu Trp Arg Arg Phe Arg Trp Ala Ile Ile Leu Phe
1 5 10 15
Ile Ile Leu Phe Ile Leu Leu Leu Phe Leu Ala Ile Phe Ile Tyr Ala
20 25 30
Phe Pro Asn Tyr Ala Ala Met Lys Leu Val Lys Pro Phe Ser Xaa
35 40 45

<210> 123
<211> 108
<212> PRT
<213> Homo sapiens

<400> 123
Met His Gln Asp Trp Leu Cys Asn Leu Gly Trp Pro Leu Leu Ser Leu
1 5 10 15
Trp Ala Ala Glu Ser Ala Pro His Val Ala Met Ala Ser Ala Thr Ala
20 25 30
Gln Leu Trp Ser Arg Pro Cys Gly Arg Thr His Met Val Ser Leu Ala
35 40 45
Leu Gly His Gln Glu Thr Gly Leu Trp Leu Cys Ser Ala Phe Gly Cys
50 55 60
Val Val Asp Ser Pro Trp Ala Ser Val Cys Pro Ser Val Lys Gly Gln
65 70 75 80
Leu Thr Val Cys Gly Ile Leu Pro Arg Val Pro Val Cys Val Tyr Val
85 90 95
Cys Ala Cys Val Arg Val Ser Met Cys Val His Ile
100 105

<210> 124
<211> 60
<212> PRT
<213> Homo sapiens

82

<400> 124

Met Arg Gly Cys Val Pro Ala Phe Leu Leu His Val Leu Ser Leu Arg
1 5 10 15
Arg Ala Cys Cys Thr Gln Ala Ala Gln Val Phe Thr Ala Gln Leu Pro
20 25 30
Gly Arg Gln Val Ala Arg Arg Arg Gly Gly Trp His Glu Gln Gln Gly
35 40 45
Gly Pro Met Leu Cys Ser Ser His His Ser Arg Thr
50 55 60

<210> 125

<211> 248

<212> PRT

<213> Homo sapiens

<400> 125

Met Ala Met Leu Pro Leu Val Leu His Trp Phe Phe Ile Glu Trp Tyr
1 5 10 15
Ser Gly Lys Lys Ser Ser Ser Ala Leu Phe Gln His Ile Thr Ala Leu
20 25 30
Phe Glu Cys Ser Met Ala Ala Ile Ile Thr Leu Leu Val Ser Asp Pro
35 40 45
Val Gly Val Leu Tyr Ile Arg Ser Cys Arg Val Leu Met Leu Ser Asp
50 55 60
Trp Tyr Thr Met Leu Tyr Asn Pro Ser Pro Asp Tyr Val Thr Thr Val
65 70 75 80
His Cys Thr His Glu Ala Val Tyr Pro Leu Tyr Thr Ile Val Phe Ile
85 90 95
Tyr Tyr Ala Phe Cys Leu Val Leu Met Met Leu Leu Arg Pro Leu Leu
100 105 110
Val Lys Lys Ile Ala Cys Gly Leu Gly Lys Ser Asp Arg Phe Lys Ser
115 120 125
Ile Tyr Ala Ala Leu Tyr Phe Phe Pro Ile Leu Thr Val Leu Gln Ala
130 135 140
Val Gly Gly Gly Leu Leu Tyr Tyr Ala Phe Pro Tyr Ile Ile Leu Val
145 150 155 160
Leu Ser Leu Val Thr Leu Ala Val Tyr Met Ser Ala Ser Glu Ile Glu
165 170 175
Asn Cys Tyr Asp Leu Leu Val Arg Lys Lys Arg Leu Ile Val Leu Phe
180 185 190
Ser His Trp Leu Leu His Ala Tyr Gly Ile Ile Ser Ile Ser Arg Val

205

Tyr Cys Gly Tyr Ile Asn Arg Leu Tyr Val Gln Tyr Tyr His Cys Thr

84

210 215 220
 Tyr Lys Gln Arg Met Ile Cys Glu Lys Met Ala Asn Pro Val Gln Leu
 225 230 235 240

Gly Ser Thr Tyr Phe Arg Glu Ala
 245

<210> 127
 <211> 612
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (245)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (246)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (249)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 127
 Met Ala Ala Ala Gly Arg Leu Pro Ser Ser Trp Ala Leu Phe Ser Pro
 1 5 10 15

Leu Leu Ala Gly Leu Ala Leu Leu Gly Val Gly Pro Val Pro Ala Arg
 20 25 30

Ala Leu His Asn Val Thr Ala Glu Leu Phe Gly Ala Glu Ala Trp Gly
 35 40 45

Thr Leu Ala Ala Phe Gly Asp Leu Asn Ser Asp Lys Gln Thr Asp Leu
 50 55 60

Phe Val Leu Arg Glu Arg Asn Asp Leu Ile Val Phe Leu Ala Asp Gln
 65 70 75 80

Asn Ala Pro Tyr Phe Lys Pro Lys Val Lys Val Ser Phe Lys Asn His
 85 90 95

Ser Ala Leu Ile Thr Ser Val Val Pro Gly Asp Tyr Asp Gly Asp Ser
 100 105 110

Gln Met Asp Val Leu Leu Thr Tyr Leu Pro Lys Asn Tyr Ala Lys Ser
 115 120 125

Glu Leu Gly Ala Val Ile Phe Trp Gly Gln Asn Gln Thr Leu Asp Pro
 130 135 140

Asn Asn Met Thr Ile Leu Asn Arg Thr Phe Gln Asp Glu Pro Leu Ile

85

145 150 155 160
 Met Asp Phe Asn Gly Asp Leu Ile Pro Asp Ile Phe Gly Ile Thr Asn
 165 170 175
 Glu Ser Asn Gln Pro Gln Ile Leu Leu Gly Gly Asn Leu Ser Trp His
 180 185 190
 Pro Ala Leu Thr Thr Thr Ser Lys Met Arg Ile Pro His Ser His Ala
 195 200 205
 Phe Ile Asp Leu Thr Glu Asp Phe Thr Ala Asp Leu Phe Leu Thr Thr
 210 215 220
 Leu Asn Ala Thr Thr Ser Thr Phe Gln Phe Glu Ile Trp Glu Asn Leu
 225 230 235 240
 Asp Gly Asn Phe Xaa Xaa Ser Thr Xaa Leu Glu Lys Pro Gln Asn Met
 245 250 255
 Met Val Val Gly Gln Ser Ala Phe Ala Asp Phe Asp Gly Asp Gly His
 260 265 270
 Met Asp His Leu Leu Pro Gly Cys Glu Asp Lys Asn Cys Gln Lys Ser
 275 280 285
 Thr Ile Tyr Leu Val Arg Ser Gly Met Lys Gln Trp Val Pro Val Leu
 290 295 300
 Gln Asp Phe Ser Asn Lys Gly Thr Leu Trp Gly Phe Val Pro Phe Val
 305 310 315 320
 Asp Glu Gln Gln Pro Thr Glu Ile Pro Ile Pro Ile Thr Leu His Ile
 325 330 335
 Gly Asp Tyr Asn Met Asp Gly Tyr Pro Asp Ala Leu Val Ile Leu Lys
 340 345 350
 Asn Thr Ser Gly Ser Asn Gln Gln Ala Phe Leu Leu Glu Asn Val Pro
 355 360 365
 Cys Asn Asn Ala Ser Cys Glu Glu Ala Arg Arg Met Phe Lys Val Tyr
 370 375 380
 Trp Glu Leu Thr Asp Leu Asn Gln Ile Lys Asp Ala Met Val Ala Thr
 385 390 395 400
 Phe Phe Asp Ile Tyr Glu Asp Gly Ile Leu Asp Ile Val Val Leu Ser
 405 410 415
 Lys Gly Tyr Thr Lys Asn Asp Phe Ala Ile His Thr Leu Lys Asn Asn
 420 425 430
 Phe Glu Ala Asp Ala Tyr Phe Val Lys Val Ile Val Leu Ser Gly Leu
 435 440 445
 Cys Ser Asn Asp Cys Pro Arg Lys Ile Thr Pro Phe Gly Val Asn Gln
 450 455 460

86

Pro Gly Pro Tyr Ile Met Tyr Thr Thr Val Asp Ala Asn Gly Tyr Leu
465 470 475 480

Lys Asn Gly Ser Ala Gly Gln Leu Ser Gln Ser Ala His Leu Ala Leu
485 490 495

Gln Leu Pro Tyr Asn Val Leu Gly Leu Gly Arg Ser Ala Asn Phe Leu
500 505 510

Asp His Leu Tyr Val Gly Ile Pro Arg Pro Ser Gly Glu Lys Ser Ile
515 520 525

Arg Lys Gln Glu Trp Thr Ala Ile Ile Pro Asn Ser Gln Leu Ile Val
530 535 540

Ile Pro Tyr Pro His Asn Val Pro Arg Ser Trp Ser Ala Lys Leu Tyr
545 550 555 560

Leu Thr Pro Ser Asn Ile Val Leu Leu Thr Ala Ile Ala Leu Ile Gly
565 570 575

Val Cys Val Phe Ile Leu Ala Ile Ile Gly Ile Leu His Trp Gln Glu
580 585 590

Lys Lys Ala Asp Asp Arg Glu Lys Arg Gln Glu Ala His Arg Phe His
595 600 605

Phe Asp Ala Met
610

<210> 128

<211> 447

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (8)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (309)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (333)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 128

87

Met Glu Leu Ser Gln Met Ser Xaa Leu Met Gly Leu Ser Val Leu Leu
 1 5 10 15
 Gly Leu Leu Ala Leu Met Ala Thr Ala Ala Val Xaa Arg Gly Trp Leu
 20 25 30
 Arg Ala Gly Glu Glu Arg Ser Gly Arg Pro Ala Cys Gln Lys Ala Asn
 35 40 45
 Gly Phe Pro Pro Asp Lys Ser Ser Gly Ser Lys Lys Gln Lys Gln Tyr
 50 55 60
 Gln Arg Ile Arg Lys Glu Lys Pro Gln Gln His Asn Phe Thr His Arg
 65 70 75 80
 Leu Leu Ala Ala Ala Leu Lys Ser His Ser Gly Asn Ile Ser Cys Met
 85 90 95
 Asp Phe Ser Ser Asn Gly Lys Tyr Leu Ala Thr Cys Ala Asp Asp Arg
 100 105 110
 Thr Ile Arg Ile Trp Ser Thr Lys Asp Phe Leu Gln Arg Glu His Arg
 115 120 125
 Ser Met Arg Ala Asn Val Glu Leu Asp His Ala Thr Leu Val Arg Phe
 130 135 140
 Ser Pro Asp Cys Arg Ala Phe Ile Val Trp Leu Ala Asn Gly Asp Thr
 145 150 155 160
 Leu Arg Val Phe Lys Met Thr Lys Arg Glu Asp Gly Gly Tyr Thr Phe
 165 170 175
 Thr Ala Thr Pro Glu Asp Phe Pro Lys Lys His Lys Ala Pro Val Ile
 180 185 190
 Asp Ile Gly Ile Ala Asn Thr Gly Lys Phe Ile Met Thr Ala Ser Ser
 195 200 205
 Asp Thr Thr Val Leu Ile Trp Ser Leu Lys Gly Gln Val Leu Ser Thr
 210 215 220
 Ile Asn Thr Asn Gln Met Asn Asn Thr His Ala Ala Val Ser Pro Cys
 225 230 235 240
 Gly Arg Phe Val Ala Ser Cys Gly Phe Thr Pro Asp Val Lys Val Trp
 245 250 255
 Glu Val Cys Phe Gly Lys Lys Gly Glu Phe Gln Glu Val Val Arg Ala
 260 265 270
 Phe Glu Leu Lys Gly His Ser Ala Ala Val His Ser Phe Ala Phe Ser
 275 280 285
 Asn Asp Ser Arg Arg Met Ala Ser Val Ser Lys Asp Gly Thr Trp Lys
 290 295 300
 Leu Trp Asp Thr Xaa Val Glu Tyr Lys Lys Lys Gln Asp Pro Tyr Leu

88

305 310 315 320
 Leu Lys Thr Gly Arg Phe Glu Glu Ala Ala Gly Ala Xaa Pro Cys Arg
 325 330 335
 Leu Ala Leu Ser Pro Asn Ala Gln Val Leu Ala Leu Ala Ser Gly Ser
 340 345 350
 Ser Ile His Leu Tyr Asn Thr Arg Arg Gly Glu Lys Glu Glu Cys Phe
 355 360 365
 Glu Arg Val His Gly Glu Cys Ile Ala Asn Leu Ser Phe Asp Ile Thr
 370 375 380
 Gly Arg Phe Leu Ala Ser Cys Gly Asp Arg Ala Val Arg Leu Phe His
 385 390 395 400
 Asn Thr Pro Gly His Arg Ala Met Val Glu Glu Met Gln Gly His Leu
 405 410 415
 Lys Arg Ala Ser Asn Glu Ser Thr Arg Gln Arg Leu Gln Gln Gln Leu
 420 425 430
 Thr Gln Ala Gln Glu Thr Leu Lys Ser Leu Gly Ala Leu Lys Lys
 435 440 445

<210> 129

<211> 291

<212> PRT

<213> Homo sapiens

<400> 129

Met Leu Phe Leu Phe Ser Met Ala Thr Leu Leu Arg Thr Ser Phe Ser
 1 5 10 15
 Asp Pro Gly Val Ile Pro Arg Ala Leu Pro Asp Glu Ala Ala Phe Ile
 20 25 30
 Glu Met Glu Ile Glu Ala Thr Asn Gly Ala Val Pro Gln Gly Gln Arg
 35 40 45
 Pro Pro Pro Arg Ile Lys Asn Phe Gln Ile Asn Asn Gln Ile Val Lys
 50 55 60
 Leu Lys Tyr Cys Tyr Thr Cys Lys Ile Phe Arg Pro Pro Arg Ala Ser
 65 70 75 80
 His Cys Ser Ile Cys Asp Asn Cys Val Glu Arg Phe Asp His His Cys
 85 90 95
 Pro Trp Val Gly Asn Cys Val Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr
 100 105 110
 Leu Phe Ile Leu Ser Leu Ser Leu Leu Thr Ile Tyr Val Phe Ala Phe
 115 120 125
 Asn Ile Val Tyr Val Ala Leu Lys Ser Leu Lys Ile Gly Phe Leu Glu

89

130 135 140
 Thr Leu Lys Glu Thr Pro Gly Thr Val Leu Glu Val Leu Ile Cys Phe
 145 150 155 160
 Phe Thr Leu Trp Ser Val Val Gly Leu Thr Gly Phe His Thr Phe Leu
 165 170 175
 Val Ala Leu Asn Gln Thr Thr Asn Glu Asp Ile Lys Gly Ser Trp Thr
 180 185 190
 Gly Lys Asn Arg Val Gln Asn Pro Tyr Ser His Gly Asn Ile Val Lys
 195 200 205
 Asn Cys Cys Glu Val Leu Cys Gly Pro Leu Pro Pro Ser Val Leu Asp
 210 215 220
 Arg Arg Gly Ile Leu Pro Leu Glu Glu Ser Gly Ser Arg Pro Pro Ser
 225 230 235 240
 Thr Gln Glu Thr Ser Ser Ser Leu Leu Pro Gln Ser Pro Ala Pro Thr
 245 250 255
 Glu His Leu Asn Ser Asn Glu Met Pro Glu Asp Ser Ser Thr Pro Glu
 260 265 270
 Glu Met Pro Pro Pro Glu Pro Pro Glu Pro Pro Gln Glu Ala Ala Glu
 275 280 285
 Ala Glu Lys
 290

<210> 130
 <211> 78
 <212> PRT
 <213> Homo sapiens

<400> 130
 Met Val Arg Lys Trp Leu Thr Phe Val Glu His Leu Leu Cys Ala Trp
 1 5 10 15
 Pro Arg Leu Gly Ala Phe Val Pro Arg Val Thr Pro Ser Glu Cys Ser
 20 25 30
 Ser Leu Pro His Ser Asn Trp Gly Val Gly Gly Arg Ala Ala Gln Leu
 35 40 45
 Thr Gly Ala Glu Leu Lys Thr His Ser Trp Val Cys Leu Gly Trp Ala
 50 55 60
 Val Leu Val Ala Pro Val Ala Asn Thr Arg Ala Pro Phe Thr
 65 70 75

<210> 131
 <211> 333
 <212> PRT

90

<213> Homo sapiens

<220>

<221> SITE

<222> (97)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 131

```

Met Leu Met Phe Ala Val Ile Val Ala Ser Ser Gly Leu Leu Leu Met
 1             5             10             15

Ile Glu Arg Gly Ile Leu Ala Glu Met Lys Pro Leu Pro Leu His Pro
 20             25             30

Pro Gly Arg Glu Gly Thr Ala Trp Arg Gly Lys Ala Pro Lys Pro Gly
 35             40             45

Gly Leu Ser Leu Arg Ala Gly Asp Ala Asp Leu Gln Val Arg Gln Asp
 50             55             60

Val Arg Asn Arg Thr Leu Arg Ala Val Cys Gly Gln Pro Gly Met Pro
 65             70             75             80

Arg Asp Pro Trp Asp Leu Pro Val Gly Gln Arg Arg Thr Leu Leu Arg
 85             90             95

Xaa Ile Leu Val Ser Asp Arg Tyr Arg Phe Leu Tyr Cys Tyr Val Pro
 100            105            110

Lys Val Ala Cys Ser Asn Trp Lys Arg Val Met Lys Val Leu Ala Gly
 115            120            125

Val Leu Asp Ser Val Asp Val Arg Leu Lys Met Asp His Arg Ser Asp
 130            135            140

Leu Val Phe Leu Ala Asp Leu Arg Pro Glu Glu Ile Arg Tyr Arg Leu
 145            150            155            160

Gln His Tyr Phe Lys Phe Leu Phe Val Arg Glu Pro Leu Glu Arg Leu
 165            170            175

Leu Ser Ala Tyr Arg Asn Lys Phe Gly Glu Ile Arg Glu Tyr Gln Gln
 180            185            190

Arg Tyr Gly Ala Glu Ile Val Arg Arg Tyr Arg Ala Gly Ala Gly Pro
 195            200            205

Ser Pro Ala Gly Asp Asp Val Thr Phe Pro Glu Phe Leu Arg Tyr Leu
 210            215            220

Val Asp Glu Asp Pro Glu Arg Met Asn Glu His Trp Met Pro Val Tyr
 225            230            235            240

His Leu Cys Gln Pro Cys Ala Val His Tyr Asp Phe Val Gly Ser Tyr
 245            250            255

Glu Arg Leu Glu Ala Asp Ala Asn Gln Val Leu Glu Trp Val Arg Ala
 260            265            270

```

91

Pro Pro His Val Arg Phe Pro Ala Arg Gln Ala Trp Tyr Arg Pro Ala
 275 280 285

Ser Pro Glu Ser Leu His Tyr His Leu Cys Ser Ala Pro Arg Ala Leu
 290 295 300

Leu Gln Asp Val Leu Pro Lys Tyr Ile Leu Asp Phe Ser Leu Phe Ala
 305 310 315 320

Tyr Pro Leu Pro Asn Val Thr Lys Glu Ala Cys Gln Gln
 325 330

<210> 132

<211> 164

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (126)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 132

Met Leu Pro Leu Leu Ile Ile Cys Leu Leu Pro Ala Ile Glu Gly Lys
 1 5 10 15

Asn Cys Leu Arg Cys Trp Pro Glu Leu Ser Ala Leu Ile Asp Tyr Asp
 20 25 30

Leu Gln Ile Leu Trp Val Thr Pro Gly Pro Pro Thr Glu Leu Ser Gln
 35 40 45

Ser Ile His Ser Leu Phe Leu Glu Asp Asn Asn Phe Leu Lys Pro Trp
 50 55 60

Tyr Leu Asp Arg Asp His Leu Glu Glu Glu Thr Ala Lys Phe Phe Thr
 65 70 75 80

Gln Val His Gln Ala Ile Lys Thr Leu Arg Asp Asp Lys Thr Val Leu
 85 90 95

Leu Glu Glu Ile Tyr Thr His Lys Asn Leu Phe Thr Glu Arg Leu Asn
 100 105 110

Lys Ile Ser Asp Gly Leu Lys Glu Lys Gly Ala Pro Pro Xaa Ser Met
 115 120 125

Asn Ala Phe Pro Ala Pro Ser Pro Thr Cys Thr Pro Glu Pro Leu Gly
 130 135 140

Ser Val Cys Leu Pro Ser Thr Ser Val Ser Leu Pro Ser His Leu Pro
 145 150 155 160

Gly Ser Leu Gln

<210> 133
 <211> 245
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (245)
 <223> Xaa equals stop translation

<400> 133
 Met Val Ala Val Gly Val Tyr Ala Arg Leu Met Lys His Ala Glu Ala
 1 5 10 15
 Ala Leu Ala Cys Leu Ala Val Asp Pro Ala Ile Leu Leu Ile Val Val
 20 25 30
 Gly Val Leu Met Phe Leu Leu Thr Phe Cys Gly Cys Ile Gly Ser Leu
 35 40 45
 Arg Glu Asn Ile Cys Leu Leu Gln Thr Phe Ser Leu Cys Leu Thr Ala
 50 55 60
 Val Phe Leu Leu Gln Leu Ala Ala Gly Ile Leu Gly Phe Val Phe Ser
 65 70 75 80
 Asp Lys Ala Arg Gly Lys Val Ser Glu Ile Ile Asn Asn Ala Ile Val
 85 90 95
 His Tyr Arg Asp Asp Leu Asp Leu Gln Asn Leu Ile Asp Phe Gly Gln
 100 105 110
 Lys Lys Phe Ser Cys Cys Gly Gly Ile Ser Tyr Lys Asp Trp Ser Gln
 115 120 125
 Asn Met Tyr Phe Asn Cys Ser Glu Asp Asn Pro Ser Arg Glu Arg Cys
 130 135 140
 Ser Val Pro Tyr Ser Cys Cys Leu Pro Thr Pro Asp Gln Ala Val Ile
 145 150 155 160
 Asn Thr Met Cys Gly Gln Gly Met Gln Ala Phe Asp Tyr Leu Glu Ala
 165 170 175
 Ser Lys Val Ile Tyr Thr Asn Gly Cys Ile Asp Lys Leu Val Asn Trp
 180 185 190
 Ile His Ser Asn Leu Phe Leu Leu Gly Gly Val Ala Leu Gly Leu Ala
 195 200 205
 Ile Pro Gln Leu Val Gly Ile Leu Leu Ser Gln Ile Leu Val Asn Gln
 210 215 220
 Ile Lys Asp Gln Ile Lys Leu Gln Leu Tyr Asn Gln Gln His Arg Ala
 225 230 235 240
 Asp Pro Trp Tyr Xaa

93

245

<210> 134
 <211> 56
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (56)
 <223> Xaa equals stop translation

<400> 134
 Met Gly Thr Val Gly Leu Trp Pro Ser Trp Leu Trp Leu Pro Ala Ser
 1 5 10 15
 Trp Pro Leu Thr Ser Cys Gly Val Thr Arg Arg Arg Leu Arg Gly Pro
 20 25 30
 Gly Leu Arg Arg Thr Ser Gln Thr Gly Arg His Thr Ser Pro Cys Pro
 35 40 45
 Thr Ala Thr Trp Ala Glu Ser Xaa
 50 55

<210> 135
 <211> 55
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (47)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (51)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (55)
 <223> Xaa equals stop translation

<400> 135
 Met Ser Ile Val Met Ser Pro Leu Leu Leu Pro Ile Cys Tyr Leu Asn
 1 5 10 15
 Leu Leu Leu Phe Phe Val Asn Leu Ala Lys Asn Leu Ser Ile Leu Phe
 20 25 30
 Val Ser Ser Lys Lys Tyr Thr Phe Val Phe Met Ile Ser Leu Xaa Phe
 35 40 45
 Phe His Xaa Tyr Phe Ile Xaa

94

50

55

<210> 136
 <211> 89
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (89)
 <223> Xaa equals stop translation

<400> 136
 Met Ala Ile Ile Ser Phe Glu Leu Leu Phe Leu Met Asn Leu Pro Thr
 1 5 10 15
 Val Asn Ser Ser Asn Phe Lys Leu Ile Ile Pro Glu Asp Val Thr Leu
 20 25 30
 Ser Phe Val Ser His Leu Asp Ile Thr Val Asn His Phe Val Phe Leu
 35 40 45
 Ser Thr Phe Glu Leu Ala Gly Val Ile Glu Gly Lys Pro Leu Pro Asp
 50 55 60
 Ser Lys Ser Asp Leu Cys Pro Ile Leu Gly Gln Leu Trp Phe His Ile
 65 70 75 80
 Leu Leu Phe Phe Ile Phe Trp Val Xaa
 85

<210> 137
 <211> 62
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (62)
 <223> Xaa equals stop translation

<400> 137
 Met Arg Leu Pro Ile Ala Pro His Leu Gln Tyr Phe Met Trp Ser Val
 1 5 10 15
 Leu Leu Phe Leu Val Ile Leu Val Asp Met Lys Trp His Leu Ser Val
 20 25 30
 Ala Phe His Tyr Ile Ser Leu Met Thr Asn Gly Ile Leu Ser Pro Phe
 35 40 45
 Gln Cys Leu Leu Ala Ile His Val Ser Leu Phe Phe Val Xaa
 50 55 60

<210> 138

95

<211> 106
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (106)
 <223> Xaa equals stop translation

<400> 138
 Met Cys Leu Leu Pro Gly Gly Val Leu Leu Ile Trp Ser Cys Ala Ser
 1 5 10 15
 Gly Thr Pro Ala Ser His Thr Lys Asp Trp Gly Arg Cys Lys Phe Ser
 20 25 30
 Ala Ala Thr Lys Arg Thr Ala Glu Ser Asn Leu Glu Ser Thr Gln Leu
 35 40 45
 Met Leu Ala Ser Gln Ile Asp Pro Leu Leu Ala Glu Cys Trp His Leu
 50 55 60
 Cys Ala Ser Val Ser Ser Ser Val Asn Gly Gly Asp Lys Lys Cys Val
 65 70 75 80
 His Thr Ser Arg Ala Val Gly Arg Ile Lys Leu Cys Ser Asp Thr Ile
 85 90 95
 Arg Ala Cys Ser Gly Trp Tyr Leu Gln Xaa
 100 105

<210> 139
 <211> 52
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (52)
 <223> Xaa equals stop translation

<400> 139
 Met Ser His Ser Val Phe Ala His Tyr Ile Phe Asn Ile Leu Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Leu Leu Ile Gly Phe Leu Tyr Ser Met Pro Phe Ile
 20 25 30
 Tyr Lys Asp Thr Lys Lys Thr His Val Cys Asn Phe Asn Asn Ile Phe
 35 40 45
 Pro Ile Leu Xaa
 50

<210> 140
 <211> 119

96

<212> PRT

<213> Homo sapiens

<400> 140

Met Lys Trp Arg Arg Lys Ser Ala Tyr Trp Lys Ala Leu Lys Val Phe
1 5 10 15

Lys Leu Pro Val Glu Phe Leu Leu Leu Leu Thr Val Pro Val Val Asp
20 25 30

Pro Asp Lys Asp Asp Gln Asn Trp Lys Arg Pro Leu Asn Cys Leu His
35 40 45

Leu Val Ile Ser Pro Leu Val Val Val Leu Thr Leu Gln Ser Gly Thr
50 55 60

Tyr Gly Val Tyr Glu Ile Gly Gly Leu Val Pro Val Trp Val Val Val
65 70 75 80

Val Ile Ala Gly Thr Ala Leu Ala Ser Val Thr Phe Phe Ala Thr Ser
85 90 95

Asp Ser Gln Pro Pro Arg Leu His Trp Leu Phe Ala Phe Leu Gly Phe
100 105 110

Leu Thr Ser Ala Leu Trp Ile
115

<210> 141

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals stop translation

<400> 141

Met Cys Ser Gly Ser Phe Lys Glu Leu Tyr Leu Val Pro Ile Ser Leu
1 5 10 15

Phe Ser Thr Cys Val Leu Gly Phe Tyr Phe His Asn Phe Leu Leu Leu
20 25 30

Ile Ile Leu Phe Ser Ile Leu Leu Arg Lys Ile Thr Gly Lys Leu Phe
35 40 45

Phe Thr Tyr Tyr His Phe Ser Cys Gly Val Xaa
50 55

<210> 142

<211> 100

<212> PRT

<213> Homo sapiens

97

<220>
 <221> SITE
 <222> (100)
 <223> Xaa equals stop translation

 <400> 142
 Met Leu Phe Phe Leu Ser Leu Phe Leu Ser Leu Leu Leu Thr Leu Ser
 1 5 10 15
 Leu Pro Ser Phe Leu Pro Phe Ser Phe Phe Phe Ser Leu Phe Pro
 20 25 30
 His Leu Ser Ala Cys Leu Leu Pro Ser Leu Pro Ser Pro Pro Phe Pro
 35 40 45
 Leu Pro Pro Ser Leu Pro Ser Phe Leu Pro Ser Phe Leu Pro Ser Phe
 50 55 60
 Leu Pro Ser Leu Leu Ser Pro Ser Phe Pro Ala Phe Phe Pro Ser Phe
 65 70 75 80
 Cys Gln Leu Ala Arg Arg Ser Pro Arg Lys Ser Thr Gln Met Leu Gln
 85 90 95
 Ser Thr Ser Xaa
 100

<210> 143
 <211> 65
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (61)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (65)
 <223> Xaa equals stop translation

<400> 143
 Met Ala Val Leu Leu Ile Thr Ile Leu Leu Phe Leu Cys Leu Gly Tyr
 1 5 10 15
 Tyr Arg Val Ile Thr Glu Ile Ser Arg Lys Thr Pro Ala Cys Arg Met
 20 25 30
 Phe Thr Ser Ser Leu Ser Ser Trp Tyr Ile Met Arg Lys Leu Tyr Asp
 35 40 45
 Thr Pro Gly Glu Val Phe Leu Ser His Ala Ile Val Xaa Phe Leu Lys
 50 55 60
 Xaa
 65

<210> 144
<211> 67
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (67)
<223> Xaa equals stop translation

<400> 144
Met Leu Asn Gln Pro Cys Ile Leu Gly Met Lys Pro Thr Trp Leu Trp
1 5 10 15
Trp Ile Ser Phe Leu Met Cys Cys Trp Val Trp Leu Ala Ser Val Leu
20 25 30
Leu Gly Ile Phe Ala Ser Ile Phe Ile Arg Asp Ile Gly Leu Glu Phe
35 40 45
Ser Phe Phe Val Met Cys Leu Pro Gly Phe Gly Ile Arg Val Met Leu
50 55 60
Ala Ser Xaa
65

<210> 145
<211> 59
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (59)
<223> Xaa equals stop translation

<400> 145
Met Thr Ala Met Ser Ile His Leu Phe Cys Thr Ala Leu Ser Cys Gly
1 5 10 15
Ser Ser Gly Gln Cys Asn Lys Ala Ile Lys Arg Asn Lys Ile Ser Asn
20 25 30
Asp Trp Lys Asp Val Asn Val Ser Ser Phe Ile Glu Asn Met Ile His
35 40 45
Arg Tyr Thr Tyr Thr Asn Ala Leu Asn Ser Xaa
50 55

<210> 146
<211> 56
<212> PRT
<213> Homo sapiens

99

<220>

<221> SITE

<222> (56)

<223> Xaa equals stop translation

<400> 146

Met Ser His Cys Thr Trp Pro Val Cys Leu Phe Cys Leu Val Pro Pro
1 5 10 15

Pro Met Gly Asp Leu Lys Glu Val Cys Leu Pro His Arg Cys Pro Gly
20 25 30

Arg Thr Ala Cys Cys Ser Tyr Ser Glu Pro His Leu Gln Thr Glu Glu
35 40 45

Asp Arg Arg Thr Leu Ile Cys Xaa
50 55

<210> 147

<211> 66

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (66)

<223> Xaa equals stop translation

<400> 147

Met Thr Asn Gly His Gln Val Leu Leu Leu Leu Leu Thr Ser Ala
1 5 10 15

Val Ala Ala Gly Pro Trp Pro Gln Val His Ala Gly Gln Trp Gly Trp
20 25 30

Met Cys Leu Pro Pro Gly Leu Pro Ser Val Gln Ala Arg Ser Gly Leu
35 40 45

Gly Gly Leu Pro Gly Gly Pro Gln Trp Val Pro Gly Gly Ala Arg Gly
50 55 60

Tyr Xaa
65

<210> 148

<211> 328

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (328)

<223> Xaa equals stop translation

<400> 148

Met Ala Cys Arg Lys Leu Ala Val Ala His Pro Leu Leu Leu Leu Arg

										100									
1		5				10				15									
His	Leu	Pro	Met	Ile	Ala	Ala	Leu	Leu	His	Gly	Arg	Thr	His	Leu	Asn				
			20					25					30						
Phe	Gln	Glu	Phe	Arg	Gln	Gln	Asn	His	Leu	Ser	Cys	Phe	Leu	His	Val				
		35					40					45							
Leu	Gly	Leu	Leu	Glu	Leu	Leu	Gln	Pro	His	Val	Phe	Arg	Ser	Glu	His				
	50					55					60								
Gln	Gly	Ala	Leu	Trp	Asp	Cys	Leu	Leu	Ser	Phe	Ile	Arg	Leu	Leu	Leu				
	65				70					75					80				
Asn	Tyr	Arg	Lys	Ser	Ser	Arg	His	Leu	Ala	Ala	Phe	Ile	Asn	Lys	Phe				
				85					90					95					
Val	Gln	Phe	Ile	His	Lys	Tyr	Ile	Thr	Tyr	Asn	Ala	Pro	Ala	Ala	Ile				
			100					105					110						
Ser	Phe	Leu	Gln	Lys	His	Ala	Asp	Pro	Leu	His	Asp	Leu	Ser	Phe	Asp				
		115					120					125							
Asn	Ser	Asp	Leu	Val	Met	Leu	Lys	Ser	Leu	Leu	Ala	Gly	Leu	Ser	Leu				
		130					135					140							
Pro	Ser	Arg	Asp	Asp	Arg	Thr	Asp	Arg	Gly	Leu	Asp	Glu	Glu	Gly	Glu				
	145				150					155					160				
Glu	Glu	Ser	Ser	Ala	Gly	Ser	Leu	Pro	Leu	Val	Ser	Val	Ser	Leu	Phe				
				165					170					175					
Thr	Pro	Leu	Thr	Ala	Ala	Glu	Met	Ala	Pro	Tyr	Met	Lys	Arg	Leu	Ser				
			180					185					190						
Arg	Gly	Gln	Thr	Val	Glu	Asp	Leu	Leu	Glu	Val	Leu	Ser	Asp	Ile	Asp				
		195					200					205							
Glu	Met	Ser	Arg	Arg	Arg	Pro	Glu	Ile	Leu	Ser	Phe	Phe	Ser	Thr	Asn				
	210					215					220								
Leu	Gln	Arg	Leu	Met	Ser	Ser	Ala	Glu	Glu	Cys	Cys	Arg	Asn	Leu	Ala				
	225				230					235				240					
Phe	Ser	Leu	Ala	Leu	Arg	Ser	Met	Gln	Asn	Ser	Pro	Ser	Ile	Ala	Ala				
				245					250					255					
Ala	Phe	Leu	Pro	Thr	Phe	Met	Tyr	Cys	Leu	Gly	Ser	Gln	Asp	Phe	Glu				
			260					265					270						
Val	Val	Gln	Thr	Ala	Leu	Arg	Asn	Leu	Pro	Glu	Tyr	Ala	Leu	Leu	Cys				
		275					280					285							
Gln	Glu	His	Ala	Ala	Val	Leu	Leu	His	Arg	Ala	Phe	Leu	Val	Gly	Met				
	290					295					300								
Tyr	Gly	Gln	Met	Asp	Pro	Ser	Ala	Gln	Ile	Ser	Glu	Ala	Leu	Arg	Ile				
	305					310				315				320					

101

Leu His Met Glu Ala Val Met Xaa
325

<210> 149
<211> 90
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (10)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (13)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (90)
<223> Xaa equals stop translation

<400> 149
Met Gly Phe Leu Gln Leu Leu Val Val Xaa Val Leu Xaa Ser Glu His
1 5 10 15
Arg Val Ala Gly Ala Ala Glu Val Phe Gly Asn Ser Ser Glu Gly Leu
20 25 30
Ile Glu Phe Ser Val Gly Lys Phe Arg Tyr Phe Glu Leu Asn Arg Pro
35 40 45
Phe Pro Glu Glu Ala Ile Leu His Asp Ile Ser Ser Asn Val Thr Phe
50 55 60
Leu Ile Phe Gln Ile His Ser Gln Tyr^AGln Asn Thr Thr Val Ser Phe
65 70 75 80
Ser Pro Arg Arg Arg Ser Pro Thr Met Xaa
85 90

<210> 150
<211> 149
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (149)
<223> Xaa equals stop translation

<400> 150
Met Ala Gly Ser Pro Leu Leu Trp Gly Pro Arg Ala Gly Gly Val Gly
1 5 10 15

102

Leu Leu Val Leu Leu Leu Leu Gly Leu Phe Arg Pro Pro Pro Ala Leu
 20 25 30
 Cys Ala Arg Pro Val Lys Glu Pro Arg Gly Leu Ser Ala Ala Ser Pro
 35 40 45
 Pro Leu Ala Arg Leu Ala Leu Leu Ala Ala Ser Gly Gly Gln Cys Pro
 50 55 60
 Glu Val Arg Arg Arg Gly Arg Cys Arg Pro Gly Ala Gly Ala Gly Ala
 65 70 75 80
 Ser Ala Gly Ala Glu Arg Gln Glu Arg Ala Arg Ala Glu Ala Gln Arg
 85 90 95
 Leu Arg Ile Ser Arg Arg Ala Ser Trp Arg Ser Cys Cys Ala Ser Gly
 100 105 110
 Ala Pro Pro Ala Thr Leu Ile Arg Leu Trp Ala Trp Thr Thr Thr Pro
 115 120 125
 Thr Arg Leu Gln Arg Ser Ser Leu Ala Leu Cys Ser Ala Pro Ala Leu
 130 135 140
 Thr Leu Pro Pro Xaa
 145

<210> 151
 <211> 391
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (391)
 <223> Xaa equals stop translation

<400> 151
 Met Leu Pro Thr Trp Trp Ile Val Ser Ser Trp Leu Val Trp Gly Val
 1 5 10 15
 Ile Leu Phe Val Tyr Leu Val Ile Arg Ala Leu Arg Leu Trp Arg Thr
 20 25 30
 Ala Lys Leu Gln Val Thr Leu Lys Lys Tyr Ser Val His Leu Glu Asp
 35 40 45
 Met Ala Thr Asn Ser Arg Ala Phe Thr Asn Leu Val Arg Lys Ala Leu
 50 55 60
 Arg Leu Ile Gln Glu Thr Glu Val Ile Ser Arg Gly Phe Thr Leu Val
 65 70 75 80
 Ser Ala Ala Cys Pro Phe Asn Lys Ala Gly Gln His Pro Ser Gln His
 85 90 95

103

Leu Ile Gly Leu Arg Lys Ala Val Tyr Arg Thr Leu Arg Ala Asn Phe
 100 105 110
 Gln Ala Ala Arg Leu Ala Thr Leu Tyr Met Leu Lys Asn Tyr Pro Leu
 115 120 125
 Asn Ser Glu Ser Asp Asn Val Thr Asn Tyr Ile Cys Val Val Pro Phe
 130 135 140
 Lys Glu Leu Gly Leu Gly Leu Ser Glu Glu Gln Ile Ser Glu Glu Glu
 145 150 155 160
 Ala His Asn Phe Thr Asp Gly Phe Ser Leu Pro Ala Leu Lys Val Leu
 165 170 175
 Phe Gln Leu Trp Val Ala Gln Ser Ser Glu Phe Phe Arg Arg Leu Ala
 180 185 190
 Leu Leu Leu Ser Thr Ala Asn Ser Pro Pro Gly Pro Leu Leu Thr Pro
 195 200 205
 Ala Leu Leu Pro His Arg Ile Leu Ser Asp Val Thr Gln Gly Leu Pro
 210 215 220
 His Ala His Ser Ala Cys Leu Glu Glu Leu Lys Arg Ser Tyr Glu Phe
 225 230 235 240
 Tyr Arg Tyr Phe Glu Thr Gln His Gln Ser Val Pro Gln Cys Leu Ser
 245 250 255
 Lys Thr Gln Gln Lys Ser Arg Glu Leu Asn Asn Val His Thr Ala Val
 260 265 270
 Arg Ser Leu Gln Leu His Leu Lys Ala Leu Leu Asn Glu Val Ile Ile
 275 280 285
 Leu Glu Asp Glu Leu Glu Lys Leu Val Cys Thr Lys Glu Thr Gln Glu
 290 295 300
 Leu Val Ser Glu Ala Tyr Pro Ile Leu Glu Gln Lys Leu Lys Leu Ile
 305 310 315 320
 Gln Pro His Val Gln Ala Ser Asn Asn Cys Trp Glu Glu Ala Ile Ser
 325 330 335
 Gln Val Asp Lys Leu Leu Arg Arg Asn Thr Asp Lys Lys Gly Lys Pro
 340 345 350
 Glu Ile Ala Cys Glu Asn Pro His Cys Thr Val Ser Thr Phe Glu Ala
 355 360 365
 Ala Tyr Ser Thr His Cys Arg Gln Arg Ser Asn Pro Arg Gly Ala Gly
 370 375 380
 Ile Arg Ser Leu Cys Arg Xaa
 385 390

104

<210> 152
<211> 99
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (99)
<223> Xaa equals stop translation

<400> 152
Met Thr Thr Arg Gln Pro Thr Ala Val Ser Trp Pro Cys Trp Leu Met
1 5 10 15
Ser Ser Ser Leu Ser Thr Ala Cys Leu Ala Trp Thr Leu Thr Gly Ser
20 25 30
Leu Ala Arg Glu Ala Thr Arg Arg Ala Arg Ser Leu Ser Pro Thr Trp
35 40 45
Asn Cys Ser Ala Arg Gln Val Pro Pro Ser Pro Pro His Ser Gly Leu
50 55 60
Gly Arg Arg Gly Trp Ala His Cys His Leu Thr Cys Leu Leu Val Thr
65 70 75 80
Gln Leu Phe Arg Val Gly Arg Ile His Pro Ile Leu Ser Leu Pro Leu
85 90 95
Val Thr Xaa

<210> 153
<211> 61
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (61)
<223> Xaa equals stop translation

<400> 153
Met Ser His Cys Ala Arg Pro Thr Phe Leu Thr Leu Leu Leu Ala Ser
1 5 10 15
Cys Phe Trp Ala Ala Ala Ile Pro Asn Arg Asn Val Ile Leu Ser Val
20 25 30
Ser Phe Arg Pro Leu His Met Gln Phe Thr Leu Ser Ile Leu Val Phe
35 40 45
Ile Leu Arg Ile Leu Ile Leu Leu Arg Ser Phe Leu Xaa
50 55 60

<210> 154

105

<211> 393
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (393)
 <223> Xaa equals stop translation

<400> 154
 Met Glu Trp Trp Ala Ser Ser Pro Leu Arg Leu Trp Leu Leu Phe
 1 5 10 15
 Leu Leu Pro Ser Ala Gln Gly Arg Gln Lys Glu Ser Gly Ser Lys Trp
 20 25 30
 Lys Val Phe Ile Asp Gln Ile Asn Arg Ser Leu Glu Asn Tyr Glu Pro
 35 40 45
 Cys Ser Ser Gln Asn Cys Ser Cys Tyr His Gly Val Ile Glu Glu Asp
 50 55 60
 Leu Thr Pro Phe Arg Gly Gly Ile Ser Arg Lys Met Met Ala Glu Val
 65 70 75 80
 Val Arg Arg Lys Leu Gly Thr His Tyr Gln Ile Thr Lys Asn Arg Leu
 85 90 95
 Tyr Arg Glu Asn Asp Cys Met Phe Pro Ser Arg Cys Ser Gly Val Glu
 100 105 110
 His Phe Ile Leu Glu Val Ile Gly Arg Leu Pro Asp Met Glu Met Val
 115 120 125
 Ile Asn Val Arg Asp Tyr Pro Gln Val Pro Lys Trp Met Glu Pro Ala
 130 135 140
 Ile Pro Val Phe Ser Phe Ser Lys Thr Ser Glu Tyr His Asp Ile Met
 145 150 155 160
 Tyr Pro Ala Trp Thr Phe Trp Glu Gly Gly Pro Ala Val Trp Pro Ile
 165 170 175
 Tyr Pro Thr Gly Leu Gly Arg Trp Asp Leu Phe Arg Glu Asp Leu Val
 180 185 190
 Arg Ser Ala Ala Gln Trp Pro Trp Lys Lys Lys Asn Ser Thr Ala Tyr
 195 200 205
 Phe Arg Gly Ser Arg Thr Ser Pro Glu Arg Asp Pro Leu Ile Leu Leu
 210 215 220
 Ser Arg Lys Asn Pro Lys Leu Val Asp Ala Glu Tyr Thr Lys Asn Gln
 225 230 235 240
 Ala Trp Lys Ser Met Lys Asp Thr Leu Gly Lys Pro Ala Ala Lys Asp
 245 250 255

106

Val His Leu Val Asp His Cys Lys Tyr Lys Tyr Leu Phe Asn Phe Arg
 260 265 270

Gly Val Ala Ala Ser Phe Arg Phe Lys His Leu Phe Leu Cys Gly Ser
 275 280 285

Leu Val Phe His Val Gly Asp Glu Trp Leu Glu Phe Phe Tyr Pro Gln
 290 295 300

Leu Lys Pro Trp Val His Tyr Ile Pro Val Lys Thr Asp Leu Ser Asn
 305 310 315 320

Val Gln Glu Leu Leu Gln Phe Val Lys Ala Asn Asp Asp Val Ala Gln
 325 330 335

Glu Ile Ala Glu Arg Gly Ser Gln Phe Ile Arg Asn His Leu Gln Met
 340 345 350

Asp Asp Ile Thr Cys Tyr Trp Glu Asn Leu Leu Ser Glu Tyr Ser Lys
 355 360 365

Phe Leu Ser Tyr Asn Val Thr Arg Arg Lys Gly Tyr Asp Gln Ile Ile
 370 375 380

Pro Lys Met Leu Lys Thr Glu Leu Xaa
 385 390

<210> 155
 <211> 75
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (75)
 <223> Xaa equals stop translation

<400> 155
 Met Leu Ile Leu Phe Leu Ser Val Cys Leu Phe Val Phe Leu Leu Thr
 1 5 10 15

Val Arg Ala Leu Cys Cys Arg Ser Ala Gly Val Trp Leu Arg Ser Thr
 20 25 30

Pro Asp Pro Val Cys Leu Gly Phe Ala Arg Gly Gly Cys Arg Ile Ala
 35 40 45

Met Ile Ala Ala Cys Phe Ser Ser Gly Ser Phe Val Pro Glu Gly His
 50 55 60

Pro Pro Asp Ala Ser Gln Ser Ser Pro Val Xaa
 65 70 75

<210> 156
 <211> 82
 <212> PRT

107

<213> Homo sapiens

<220>

<221> SITE

<222> (82)

<223> Xaa equals stop translation

<400> 156

Met Trp Pro Leu Leu Ala Val Ser Pro Phe Gly Leu Val Trp Ala Ser
 1 5 10 15

Ser Gln Ser Gly Ser Leu Leu Leu Arg Ala Ser Ile Pro Arg Gln His
 20 25 30

Ser Arg Arg Ala Trp His Phe Tyr Ser Glu Val Trp Gln Ser His Ser
 35 40 45

Val Ala Ser Val Leu Leu Tyr Leu Leu Val Arg Ala Ile Thr Lys Met
 50 55 60

Cys Ile Gly Ser Lys Lys Arg Asp Ile Thr Pro Thr Thr Arg Trp Lys
 65 70 75 80

Lys Xaa

<210> 157

<211> 54

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (49)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (54)

<223> Xaa equals stop translation

<400> 157

Met Ser His His Ala Gly Leu Gly Gly Gly Ile Leu Phe Ser Leu Lys
 1 5 10 15

Ile Ser Phe Phe Ile Ala Leu Ala Val Val Gly Gly Ser Arg Gly Val
 20 25 30

Asn Asp Cys Gln Leu Gly Gly Cys Arg Val Gly Ser Cys Pro Arg Val
 35 40 45

Xaa Val Arg Val Ala Xaa
 50

<210> 158

<211> 103

108

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (103)

<223> Xaa equals stop translation

<400> 158

```

Met Thr Val Arg Arg Leu Ser Leu Leu Cys Arg Asp Leu Trp Ala Leu
 1             5             10             15

Trp Leu Leu Leu Lys Ala Gly Ala Val Arg Gly Ala Arg Ala Gly Pro
          20             25             30

Arg Leu Pro Gly Arg Cys Cys Gly Ala Thr Cys Gly Asp Ala Gly Arg
          35             40             45

Gly Trp Thr Phe Trp Ala Gln Pro Cys Pro Gln Lys Leu Leu Gly Gln
 50             55             60

Lys Pro Gly Ala Gly Gly Cys Arg Gly Trp Val Leu Gly Trp Val Pro
 65             70             75             80

Pro Arg Pro Glu Glu Pro Cys Ser Leu Ala Gly Lys Val Cys Thr Gly
          85             90             95

Leu Ala Arg Trp Met Val Xaa
          100

```

<210> 159

<211> 575

<212> PRT

<213> Homo sapiens

<400> 159

```

Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala
 1             5             10             15

Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys
          20             25             30

Phe Arg Gln Glu Ser His Val Glu Phe Ala Thr Gly Glu Leu Leu Thr
          35             40             45

Met Thr Gln Trp Pro Gly Val Trp Ile Pro Met Ala Ser Cys Ser Ser
          50             55             60

Thr Trp Trp Ser Met Ala Leu Ser Pro Asp Ser Leu Ala Asp Ala Asp
          65             70             75             80

Leu Gln Val Gln Asp Phe Glu Glu His Tyr Val Gln Thr Gly Pro Gly
          85             90             95

Gln Leu Phe Val Gly Ser Thr Gln Arg Phe Phe Gln Gly Gly Leu Pro
          100             105             110

```

109

Ser Phe Leu Arg Cys Asn His Ser Ile Gln Tyr Asn Ala Ala Arg Gly
115 120 125

Pro Gln Pro Gln Leu Val Gln His Leu Arg Ala Ser Ala Ile Ser Ser
130 135 140

Ala Phe Asp Pro Glu Ala Glu Ala Leu Arg Phe Gln Leu Ala Thr Ala
145 150 155 160

Leu Gln Ala Glu Glu Asn Glu Val Gly Cys Pro Glu Gly Phe Glu Leu
165 170 175

Asp Ser Gln Gly Ala Phe Cys Val Asp Val Asp Glu Cys Ala Trp Asp
180 185 190

Ala His Leu Cys Arg Glu Gly Gln Arg Cys Val Asn Leu Leu Gly Ser
195 200 205

Tyr Arg Cys Leu Pro Asp Cys Gly Pro Gly Phe Arg Val Ala Asp Gly
210 215 220

Ala Gly Cys Glu Asp Val Asp Glu Cys Leu Glu Gly Leu Asp Asp Cys
225 230 235 240

His Tyr Asn Gln Leu Cys Glu Asn Thr Pro Gly Gly His Arg Cys Ser
245 250 255

Cys Pro Arg Gly Tyr Arg Met Gln Gly Pro Ser Leu Pro Cys Leu Asp
260 265 270

Val Asn Glu Cys Leu Gln Leu Pro Lys Ala Cys Ala Tyr Gln Cys His
275 280 285

Asn Leu Gln Gly Ser Tyr Arg Cys Leu Cys Pro Pro Gly Gln Thr Leu
290 295 300

Leu Arg Asp Gly Lys Ala Cys Thr Ser Leu Glu Arg Asn Gly Gln Asn
305 310 315 320

Val Thr Thr Val Ser His Arg Gly Pro Leu Leu Pro Trp Leu Arg Pro
325 330 335

Trp Ala Ser Ile Pro Gly Thr Ser Tyr His Ala Trp Val Ser Leu Arg
340 345 350

Pro Gly Pro Met Ala Leu Ser Ser Val Gly Arg Ala Trp Cys Pro Pro
355 360 365

Gly Phe Ile Arg Gln Asn Gly Val Cys Thr Asp Leu Asp Glu Cys Arg
370 375 380

Val Arg Asn Leu Cys Gln His Ala Cys Arg Asn Thr Glu Gly Ser Tyr
385 390 395 400

Gln Cys Leu Cys Pro Ala Gly Tyr Arg Leu Leu Pro Ser Gly Lys Asn
405 410 415

Cys Gln Asp Ile Asn Glu Cys Glu Glu Glu Ser Ile Glu Cys Gly Pro

110

420

425

430

Gly Gln Met Cys Phe Asn Thr Arg Gly Ser Tyr Gln Cys Val Asp Thr
 435 440 445

Pro Cys Pro Ala Thr Tyr Arg Gln Gly Pro Ser Pro Gly Thr Cys Phe
 450 455 460

Arg Arg Cys Ser Gln Asp Cys Gly Thr Gly Gly Pro Ser Thr Leu Gln
 465 470 475 480

Tyr Arg Leu Leu Pro Leu Pro Leu Gly Val Arg Ala His His Asp Val
 485 490 495

Ala Arg Leu Thr Ala Phe Ser Glu Val Gly Val Pro Ala Asn Arg Thr
 500 505 510

Glu Leu Ser Met Leu Glu Pro Asp Pro Arg Ser Pro Phe Ala Leu Arg
 515 520 525

Pro Leu Arg Ala Gly Leu Gly Ala Val Tyr Thr Arg Arg Ala Leu Thr
 530 535 540

Arg Ala Gly Leu Tyr Arg Leu Thr Val Arg Ala Ala Ala Pro Arg His
 545 550 555 560

Gln Ser Val Phe Val Leu Leu Ile Ala Val Ser Pro Tyr Pro Tyr
 565 570 575

<210> 160

<211> 643

<212> PRT

<213> Homo sapiens

<400> 160

Met Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Leu Val
 1 5 10 15

Leu Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu
 20 25 30

Ser Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro
 35 40 45

Val Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val
 50 55 60

Met Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu
 65 70 75 80

Pro Ser Cys Thr Asp Gln Cys Phe Arg Gln Asn His Ile His Gly Thr
 85 90 95

Leu Leu Gln Val Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys His
 100 105 110

Gly Thr Asn Ser Asp Phe Gln His Glu Leu Thr Asp Ile Thr Val Cys

111

115 120 125
Thr Lys Ala Lys Leu Trp Leu Ala Lys Arg Gln Asn Pro Cys Leu Val
130 135 140
Thr Arg Ala Val Tyr Ile Asp Ile Leu Phe Leu Leu Thr Cys Cys Leu
145 150 155 160
Asn Arg Ser Ala Lys Asp Asn Gln Pro Val Leu Glu Ser Leu Gly Phe
165 170 175
Trp Glu Glu Val Arg Gly Ile Ile Ser Gly Ser Glu Leu Ile Thr Gly
180 185 190
Phe Pro Trp Ala Phe Lys Val Pro Gly Leu Pro Gln Tyr Leu Gln Ser
195 200 205
Leu Thr Arg Leu Ala Ile Ala Ala Val Trp Ala Ala Ala Lys Ser
210 215 220
Gly Glu Arg Glu Thr Asn Val Pro Ile Ser Phe Ser Gln Leu Leu Glu
225 230 235 240
Ser Ala Phe Pro Glu Val Arg Ser Leu Thr Leu Glu Ala Leu Leu Glu
245 250 255
Lys Phe Leu Ala Ala Ala Ser Gly Leu Gly Glu Lys Gly Val Pro Pro
260 265 270
Leu Leu Cys Asn Met Gly Glu Lys Phe Leu Leu Leu Ala Met Lys Glu
275 280 285
Asn His Pro Glu Cys Phe Cys Lys Ile Leu Lys Ile Leu His Cys Met
290 295 300
Asp Pro Gly Glu Trp Leu Pro Gln Thr Glu His Cys Val His Leu Thr
305 310 315 320
Pro Lys Glu Phe Leu Ile Trp Thr Met Asp Ile Ala Ser Asn Glu Arg
325 330 335
Ser Glu Ile Gln Ser Val Ala Leu Arg Leu Ala Ser Lys Val Ile Ser
340 345 350
His His Met Gln Thr Cys Val Glu Asn Arg Glu Leu Ile Ala Ala Glu
355 360 365
Leu Lys Gln Trp Val Gln Leu Val Ile Leu Ser Cys Glu Asp His Leu
370 375 380
Pro Thr Glu Ser Arg Leu Ala Val Val Glu Val Leu Thr Ser Thr Thr
385 390 395 400
Pro Leu Phe Leu Thr Asn Pro His Pro Ile Leu Glu Leu Gln Asp Thr
405 410 415
Leu Ala Leu Trp Lys Cys Val Leu Thr Leu Leu Gln Ser Glu Glu Gln
420 425 430

112

Ala Val Arg Asp Ala Ala Thr Glu Thr Val Thr Thr Ala Met Ser Gln
 435 440 445
 Glu Asn Thr Cys Gln Ser Thr Glu Phe Ala Phe Cys Gln Val Asp Ala
 450 455 460
 Ser Ile Ala Leu Ala Leu Ala Leu Ala Val Leu Cys Asp Leu Leu Gln
 465 470 475 480
 Gln Trp Asp Gln Leu Ala Pro Gly Leu Pro Ile Leu Leu Gly Trp Leu
 485 490 495
 Leu Gly Glu Ser Asp Asp Leu Val Ala Cys Val Glu Ser Met His Gln
 500 505 510
 Val Glu Glu Asp Tyr Leu Phe Glu Lys Ala Glu Val Asn Phe Trp Ala
 515 520 525
 Glu Thr Leu Ile Phe Val Lys Tyr Leu Cys Lys His Leu Phe Cys Leu
 530 535 540
 Leu Ser Lys Ser Gly Trp Arg Pro Pro Ser Pro Glu Met Leu Cys His
 545 550 555 560
 Leu Gln Arg Met Val Ser Glu Gln Cys His Leu Leu Ser Gln Phe Phe
 565 570 575
 Arg Glu Leu Pro Pro Ala Ala Glu Phe Val Lys Thr Val Glu Phe Thr
 580 585 590
 Arg Leu Arg Ile Gln Glu Glu Arg Thr Leu Ala Cys Leu Arg Leu Leu
 595 600 605
 Ala Phe Leu Glu Gly Lys Glu Gly Glu Asp Thr Leu Val Leu Ser Val
 610 615 620
 Trp Asp Ser Tyr Ala Glu Ser Arg Gln Leu Thr Leu Pro Arg Thr Glu
 625 630 635 640
 Ala Ala Cys

<210> 161
 <211> 191
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (191)
 <223> Xaa equals stop translation

<400> 161
 Met Ser Ser Gly Thr Glu Leu Leu Trp Pro Gly Ala Ala Leu Leu Val
 1 5 10 15

113

Leu Leu Gly Val Ala Ala Ser Leu Cys Val Arg Cys Ser Arg Pro Gly
 20 25 30
 Ala Lys Arg Ser Glu Lys Ile Tyr Gln Gln Arg Ser Leu Arg Glu Asp
 35 40 45
 Gln Gln Ser Phe Thr Gly Ser Arg Thr Tyr Ser Leu Val Gly Gln Ala
 50 55 60
 Trp Pro Gly Pro Leu Ala Asp Met Ala Pro Thr Arg Lys Asp Lys Leu
 65 70 75 80
 Leu Gln Phe Tyr Pro Ser Leu Glu Asp Pro Ala Ser Ser Arg Tyr Gln
 85 90 95
 Asn Phe Ser Lys Gly Ser Arg His Gly Ser Glu Glu Ala Tyr Ile Asp
 100 105 110
 Pro Ile Ala Met Glu Tyr Tyr Asn Trp Gly Arg Phe Ser Lys Pro Pro
 115 120 125
 Glu Asp Asp Asp Ala Asn Ser Tyr Glu Asn Val Leu Ile Cys Lys Gln
 130 135 140
 Lys Thr Thr Glu Thr Gly Ala Gln Gln Glu Gly Ile Gly Gly Leu Cys
 145 150 155 160
 Arg Gly Asp Leu Ser Leu Ser Leu Ala Leu Lys Thr Gly Pro Thr Ser
 165 170 175
 Gly Leu Cys Pro Ser Ala Ser Pro Glu Glu Asp Glu Gly Ile Xaa
 180 185 190

<210> 162

<211> 64

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (64)

<223> Xaa equals stop translation

<400> 162

Met Lys His Val Leu Asn Leu Tyr Leu Leu Gly Val Val Leu Thr Leu
 1 5 10 15
 Leu Ser Ile Phe Val Arg Val Met Glu Ser Leu Glu Gly Leu Leu Glu
 20 25 30
 Ser Pro Ser Pro Gly Thr Ser Trp Thr Thr Arg Ser Gln Leu Ala Asn
 35 40 45
 Thr Glu Pro Thr Lys Gly Leu Pro Asp His Pro Ser Arg Ser Met Xaa
 50 55 60

114

<210> 163
<211> 118
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (118)
<223> Xaa equals stop translation

<400> 163
Met Ile Phe Leu Thr Val Leu Pro Leu Ala Phe Leu Phe Leu His Ser
1 5 10 15
Gly Phe Tyr His Tyr Ile Ser Phe Ser Cys Leu Phe Ser Leu Ser Leu
20 25 30
Ala Leu Phe Phe Phe Leu Asp Val Ala Thr Phe Arg Arg Pro Gly Gln
35 40 45
Leu Phe Cys Glu Arg Ser Val Leu Phe Asp Met Phe His Phe Gly Phe
50 55 60
Val Ser Leu Phe Leu His Glu Trp Ile Gln Ala Lys His Phe Trp Ala
65 70 75 80
Gly Leu Phe Ile Val Leu Pro Ser Asp Val Phe Phe Ser Val His His
85 90 95
Leu Glu Ala Pro Asp Gly Ser Phe Pro Asn Ile Ala Lys Leu Ser Leu
100 105 110
Ile Ile Leu Leu Arg Xaa
115

<210> 164
<211> 43
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (43)
<223> Xaa equals stop translation

<400> 164
Met Leu Leu Gln Phe Thr Leu Trp Val Phe Gly Ala Ile His Phe Pro
1 5 10 15
Lys Cys Leu Gly Ile Lys Glu Glu Leu Leu Lys Cys Cys Leu Gln Leu
20 25 30
Pro Pro Ser Ser Thr Tyr Glu Lys Val Val Xaa
35 40

115

<210> 165
<211> 48
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (48)
<223> Xaa equals stop translation

<400> 165
Met Leu Ser Arg Arg Leu His Cys Leu Val Leu Tyr Phe Leu Leu Leu
1 5 10 15
Leu Leu Ser Phe Ile His Thr Leu Ser Val Ser His Ile Cys Ser Ser
20 25 30
Phe Ile Trp Leu Phe Pro Lys Asn Ile Glu Ser Glu Ala Thr Met Xaa
35 40 45

<210> 166
<211> 46
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (46)
<223> Xaa equals stop translation

<400> 166
Met Glu Lys Met Gly Gln Gly Leu Leu Ser Ser Thr Tyr Leu Thr Val
1 5 10 15
Leu His Leu Ile Gln Leu Val Gly Cys Gly Leu Leu Thr Glu Glu Ile
20 25 30
Lys Glu Ser Lys Tyr Leu Ile Lys Thr Leu Gly Ser Gly Xaa
35 40 45

<210> 167
<211> 207
<212> PRT
<213> Homo sapiens

<400> 167
Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn Cys Ile Phe Phe
1 5 10 15
Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile Thr Ala Ile Ser
20 25 30

116

Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile Phe Phe Pro Leu
 35 40 45
 Pro Ala Cys Leu Asn Pro Val Leu Tyr Val Phe Phe Asn Pro Lys Phe
 50 55 60
 Lys Glu Asp Trp Lys Leu Leu Lys Arg Arg Val Thr Lys Lys Ser Gly
 65 70 75 80
 Ser Val Ser Val Ser Ile Ser Ser Gln Gly Gly Cys Leu Glu Gln Asp
 85 90 95
 Phe Tyr Tyr Asp Cys Gly Met Tyr Ser His Leu Gln Gly Asn Leu Thr
 100 105 110
 Val Cys Asp Cys Cys Glu Ser Phe Leu Leu Thr Lys Pro Val Ser Cys
 115 120 125
 Lys His Leu Ile Lys Ser His Ser Cys Pro Ala Leu Ala Val Ala Ser
 130 135 140
 Cys Gln Arg Pro Glu Gly Tyr Trp Ser Asp Cys Gly Thr Gln Ser Ala
 145 150 155 160
 His Ser Asp Tyr Ala Asp Glu Glu Asp Ser Phe Val Ser Asp Ser Ser
 165 170 175
 Asp Gln Val Gln Ala Cys Gly Arg Ala Cys Phe Tyr Gln Ser Arg Gly
 180 185 190
 Phe Pro Leu Val Arg Tyr Ala Tyr Asn Leu Pro Arg Val Lys Asp
 195 200 205

<210> 168

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (51)

<223> Xaa equals stop translation

<400> 168

Met Tyr Ile Phe Glu Leu Ser Leu Tyr Leu Glu Gly Thr Ser Phe Val
 1 5 10 15
 Val Val Leu Leu Phe Leu Leu Ile Ser Val Ser Leu Asp Ser Pro Pro
 20 25 30
 Thr Thr Lys Gly Trp Asp Ser Val Leu His Ile Trp Val Pro Leu Ile
 35 40 45
 Val Gln Xaa
 50

117

<210> 169
<211> 43
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (43)
<223> Xaa equals stop translation

<400> 169
Met Ala His Pro Gly Leu Pro Lys Thr Val Pro Val Tyr Ala Val Val
1 5 10 15
Leu Ala Leu Leu Ile Met Thr Leu Pro Leu Thr Leu Thr Ile Asn Leu
20 25 30
Asp Asp Asn Leu Tyr Gly Asn Ser Ala Lys Xaa
35 40

<210> 170
<211> 56
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation

<400> 170
Met Arg Pro Trp Trp Ser Leu Leu Leu Glu Ala Cys Ala Thr Cys Ala
1 5 10 15
Gln Thr Gly Pro Thr Arg Ser Thr Ser Cys Thr Gln Glu Val Ser His
20 25 30
Ser Ser Ser Thr Ala Tyr Pro Ala Pro Met Arg Arg Arg Cys Cys Leu
35 40 45
Pro Ser Pro Arg Ser Cys Thr Xaa
50 55

<210> 171
<211> 109
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (109)
<223> Xaa equals stop translation

<400> 171
Met Ala Leu Ala Gly Ser Val Phe Val Leu Gly Gly Val Leu Val Leu

118

1	5	10	15												
Cys	Val	Glu	Arg	Asn	Gly	Glu	Gly	Glu	Met	Gly	Trp	Pro	Gln	His	Leu
			20					25					30		
Pro	Lys	Ser	Gln	Pro	Leu	Ser	Pro	Pro	Val	Ala	Val	Arg	Arg	Cys	Ser
		35					40					45			
Phe	Glu	Arg	Ser	Trp	Ile	Asp	Leu	Leu	Val	Glu	Thr	Ser	Ser	Ser	Met
	50					55					60				
Val	Thr	Cys	Arg	Gln	Gln	Val	Gly	Thr	Pro	Asn	Gly	Met	Glu	Gly	Arg
	65				70					75					80
Gly	Gly	Gly	Pro	Lys	Thr	Thr	Phe	Pro	Ile	Arg	Leu	Gln	Leu	Ser	Gly
				85					90					95	
Ala	Cys	Ala	Val	Arg	Pro	Glu	Ile	Gln	Trp	Glu	Val	Xaa			
			100					105							

<210> 172
 <211> 51
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (17)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (51)
 <223> Xaa equals stop translation

<400> 172

Met	Phe	Leu	Phe	Phe	Tyr	Leu	Ser	Leu	Ala	Val	Tyr	Ala	Gln	Arg	Gln
1				5					10				15		
Xaa	Ser	Gly	Ser	Cys	Arg	Gln	Thr	Asp	His	Arg	Trp	Lys	Ser	Arg	Gly
		20						25					30		
Ala	Arg	Arg	Cys	Phe	Leu	Glu	Pro	Arg	Asp	Pro	Gly	Ser	Val	Pro	Gly
		35						40				45			
His	Pro	Xaa													
		50													

<210> 173
 <211> 566
 <212> PRT
 <213> Homo sapiens

<400> 173

Met	Ala	Pro	Leu	Ala	Leu	His	Leu	Leu	Val	Leu	Val	Pro	Ile	Leu	Leu
1					5				10				15		

119

Ser Leu Val Ala Ser Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro
 20 25 30
 Phe Glu Lys Cys Met Gln Asp Pro Asp Tyr Glu Gln Leu Lys Val
 35 40 45
 Val Thr Trp Gly Leu Asn Arg Thr Leu Lys Pro Gln Arg Val Ile Val
 50 55 60
 Val Gly Ala Gly Val Ala Gly Leu Val Ala Ala Lys Val Leu Ser Asp
 65 70 75 80
 Ala Gly His Lys Val Thr Ile Leu Glu Ala Asp Asn Arg Ile Gly Gly
 85 90 95
 Arg Ile Phe Thr Tyr Arg Asp Gln Asn Thr Gly Trp Ile Gly Glu Leu
 100 105 110
 Gly Ala Met Arg Met Pro Ser Ser His Arg Ile Leu His Lys Leu Cys
 115 120 125
 Gln Gly Leu Gly Leu Asn Leu Thr Lys Phe Thr Gln Tyr Asp Lys Asn
 130 135 140
 Thr Trp Thr Glu Val His Glu Val Lys Leu Arg Asn Tyr Val Val Glu
 145 150 155 160
 Lys Val Pro Glu Lys Leu Gly Tyr Ala Leu Arg Pro Gln Glu Lys Gly
 165 170 175
 His Ser Pro Glu Asp Ile Tyr Gln Met Ala Leu Asn Gln Ala Leu Lys
 180 185 190
 Asp Leu Lys Ala Leu Gly Cys Arg Lys Ala Met Lys Lys Phe Glu Arg
 195 200 205
 His Thr Leu Leu Glu Tyr Leu Leu Gly Glu Gly Asn Leu Ser Arg Pro
 210 215 220
 Ala Val Gln Leu Leu Gly Asp Val Met Ser Glu Asp Gly Phe Phe Tyr
 225 230 235 240
 Leu Ser Phe Ala Glu Ala Leu Arg Ala His Ser Cys Leu Ser Asp Arg
 245 250 255
 Leu Gln Tyr Ser Arg Ile Val Gly Gly Trp Asp Leu Leu Pro Arg Ala
 260 265 270
 Leu Leu Ser Ser Leu Ser Gly Leu Val Leu Leu Asn Ala Pro Val Val
 275 280 285
 Ala Met Thr Gln Gly Pro His Asp Val His Val Gln Ile Glu Thr Ser
 290 295 300
 Pro Pro Ala Arg Asn Leu Lys Val Leu Lys Ala Asp Val Val Leu Leu
 305 310 315 320

120

Thr Ala Ser Gly Pro Ala Val Lys Arg Ile Thr Phe Ser Pro Pro Leu
 325 330 335
 Pro Arg His Met Gln Glu Ala Leu Arg Arg Leu His Tyr Val Pro Ala
 340 345 350
 Thr Lys Val Phe Leu Ser Phe Arg Arg Pro Phe Trp Arg Glu Glu His
 355 360 365
 Ile Glu Gly Gly His Ser Asn Thr Asp Arg Pro Ser Arg Met Ile Phe
 370 375 380
 Tyr Pro Pro Pro Arg Glu Gly Ala Leu Leu Leu Ala Ser Tyr Thr Trp
 385 390 395 400
 Ser Asp Ala Ala Ala Ala Phe Ala Gly Leu Ser Arg Glu Glu Ala Leu
 405 410 415
 Arg Leu Ala Leu Asp Asp Val Ala Ala Leu His Gly Pro Val Val Arg
 420 425 430
 Gln Leu Trp Asp Gly Thr Gly Val Val Lys Arg Trp Ala Glu Asp Gln
 435 440 445
 His Ser Gln Gly Gly Phe Val Val Gln Pro Pro Ala Leu Trp Gln Thr
 450 455 460
 Glu Lys Asp Asp Trp Thr Val Pro Tyr Gly Arg Ile Tyr Phe Ala Gly
 465 470 475 480
 Glu His Thr Ala Tyr Pro His Gly Trp Val Glu Thr Ala Val Lys Leu
 485 490 495
 Leu Arg Ala Ala Ile Lys Ile Asn Ser Arg Lys Gly Pro Ala Ser Asp
 500 505 510
 Thr Ala Ser Pro Glu Gly His Ala Ser Asp Met Glu Gly Gln Gly His
 515 520 525
 Val His Gly Val Ala Ser Ser Pro Ser His Asp Leu Ala Lys Glu Glu
 530 535 540
 Gly Ser His Pro Pro Val Gln Gly Gln Leu Ser Leu Gln Asn Thr Thr
 545 550 555 560
 His Thr Arg Thr Ser His
 565

<210> 174

<211> 224

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (76)

<223> Xaa equals any of the naturally occurring L-amino acids

121

<400> 174

Met Ala Arg Ala Arg Gly Ser Pro Cys Pro Pro Leu Pro Pro Gly Arg
1 5 10 15

Met Ser Trp Pro His Gly Ala Leu Leu Phe Leu Trp Leu Phe Ser Pro
20 25 30

Pro Leu Gly Ala Gly Gly Gly Gly Val Ala Val Thr Ser Ala Ala Gly
35 40 45

Gly Gly Ser Pro Pro Ala Thr Ser Cys Pro Val Ala Cys Ser Cys Ser
50 55 60

Asn Gln Ala Ser Arg Val Ile Cys Thr Arg Arg Xaa Leu Ala Glu Val
65 70 75 80

Pro Ala Ser Ile Pro Val Asn Thr Arg Tyr Leu Asn Leu Gln Glu Asn
85 90 95

Gly Ile Gln Val Ile Arg Thr Asp Thr Phe Lys His Leu Arg His Leu
100 105 110

Glu Ile Leu Gln Leu Ser Lys Asn Leu Val Arg Lys Ile Glu Val Gly
115 120 125

Ala Phe Asn Gly Leu Pro Ser Leu Asn Thr Leu Glu Leu Phe Asp Asn
130 135 140

Arg Leu Thr Thr Val Pro Thr Gln Ala Phe Glu Tyr Leu Ser Lys Leu
145 150 155 160

Arg Glu Leu Trp Leu Arg Asn Asn Pro Ile Glu Ser Ile Pro Ser Tyr
165 170 175

Ala Phe Asn Arg Val Pro Ser Leu Arg Arg Leu Asp Leu Gly Glu Leu
180 185 190

Lys Arg Leu Glu Tyr Ile Ser Glu Ala Ala Phe Glu Gly Leu Val Asn
195 200 205

Leu Arg Tyr Leu Asn Leu Gly Met Cys Asn Leu Lys Asp Ile Pro Asn
210 215 220

<210> 175

<211> 123

<212> PRT

<213> Homo sapiens

<400> 175

Met His Asp Gly Ser Lys Pro Phe Pro Arg Tyr Gly Tyr Lys Pro Ser
1 5 10 15

Pro Pro Asn Gly Cys Gly Ser Pro Leu Phe Gly Val His Leu Asn Ile

122

20	25	30
Gly Ile Pro Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg Cys Tyr		
35	40	45
Glu Thr Cys Gly Lys Ser Lys Asn Asp Cys Asp Glu Glu Phe Gln Tyr		
50	55	60
Cys Leu Ser Lys Ile Cys Arg Asp Val Gln Lys Thr Leu Gly Leu Thr		
65	70	75
Gln His Val Gln Ala Cys Glu Thr Thr Val Glu Leu Leu Phe Asp Ser		
85	90	95
Val Ile His Leu Gly Cys Lys Pro Tyr Leu Asp Ser Gln Arg Ala Ala		
100	105	110
Cys Arg Cys His Tyr Glu Glu Lys Thr Asp Leu		
115	120	

<210> 176
 <211> 60
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (60)
 <223> Xaa equals stop translation

<400> 176

Met Gly Leu Ser Val Leu Leu Pro Leu Cys Leu Leu Gly Pro Gly Arg		
1	5	10
Phe Thr Ser Gly Gln Lys Pro Leu Asp Thr Pro Gly Leu Gly Ala Ala		
20	25	30
Val Leu Ser Val Arg Lys Ala Gly Leu Lys Met Arg Ser His Leu Thr		
35	40	45
Pro Ser Val Cys Thr Val Pro Ser Pro Gly Ser Xaa		
50	55	60

<210> 177
 <211> 105
 <212> PRT
 <213> Homo sapiens

<400> 177

Met Asp Thr Val Phe Leu Ile Gln Tyr Leu Phe Leu Thr Phe Pro Arg		
1	5	10
Ile Val Phe Met Leu Gly Phe Val Val Val Leu Ser Phe Leu Leu Gly		
20	25	30
Gly Tyr Leu Leu Phe Val Leu Tyr Leu Ala Ala Thr Asn Gln Thr Thr		

123

35

40

45

Asn Glu Trp Tyr Arg Gly Asp Trp Ala Trp Cys Gln Arg Cys Pro Leu
 50 55 60

Val Ala Trp Pro Pro Ser Ala Glu Pro Gln Val His Arg Asn Ile His
 65 70 75 80

Ser His Gly Leu Arg Ser Asn Leu Gln Glu Ile Phe Leu Pro Ala Phe
 85 90 95

Pro Cys His Glu Arg Lys Lys Gln Glu
 100 105

<210> 178

<211> 88

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (88)

<223> Xaa equals stop translation

<400> 178

Met Ala Asp Pro His Val Ser Phe Leu Ser Phe Arg Gln Leu Phe Ser
 1 5 10 15

Trp Ala Ala Val Ile Leu Leu Arg Gly Ile Leu Gly Thr Val Ala Pro
 20 25 30

Pro Pro Cys Pro Cys Val Leu Asp Leu Ala Val Tyr Pro Leu His Leu
 35 40 45

Pro Val Glu Ala Pro Cys Leu Glu Val Val Phe Lys Gln Lys Asn Gly
 50 55 60

Lys Asp Asn Cys Leu Val Phe Tyr Pro Asp Pro Ile Pro Leu Arg Gly
 65 70 75 80

Ser Leu Leu Gly Pro Phe Ile Xaa
 85

<210> 179

<211> 88

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (55)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (66)

124

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (88)

<223> Xaa equals stop translation

<400> 179

Met	Ala	Asp	Pro	His	Val	Ser	Phe	Leu	Ser	Phe	Arg	Gln	Leu	Phe	Ser
1				5					10					15	

Trp	Ala	Ala	Val	Ile	Leu	Leu	Arg	Gly	Ile	Leu	Gly	Thr	Val	Ala	Pro
			20					25						30	

Pro	Pro	Cys	Pro	Cys	Val	Leu	Asp	Leu	Ala	Val	Tyr	Pro	Leu	His	Leu
		35					40						45		

Pro	Val	Glu	Ala	Pro	Cys	Xaa	Glu	Val	Val	Phe	Lys	Gln	Lys	Asn	Gly
	50					55							60		

Lys	Xaa	Asn	Cys	Leu	Val	Phe	Tyr	Pro	Asp	Pro	Ile	Pro	Leu	Arg	Gly
65					70					75				80	

Ser	Leu	Leu	Gly	Pro	Phe	Ile	Xaa
							85

<210> 180

<211> 49

<212> PRT

<213> Homo sapiens

<400> 180

Met	Asn	Leu	Leu	Gly	Met	Ile	Phe	Ser	Met	Cys	Gly	Leu	Met	Leu	Lys
1				5					10					15	

Leu	Lys	Trp	Cys	Ala	Trp	Val	Ala	Val	Tyr	Cys	Ser	Phe	Ile	Ser	Phe
			20					25						30	

Ala	Asn	Ser	Arg	Ser	Ser	Glu	Asp	Thr	Lys	Gln	Met	Met	Ser	Ser	Phe
			35				40						45		

Met

<210> 181

<211> 23

<212> PRT

<213> Homo sapiens

<400> 181

Leu	Gly	Ser	Leu	Ser	Thr	Ala	Pro	Ser	Ser	Ala	Leu	Pro	Thr	Leu	Gly
1				5					10					15	

Ala	Arg	Arg	Thr	Arg	Ser	Lys
						20

125

<210> 182
 <211> 104
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (104)
 <223> Xaa equals stop translation

<400> 182
 Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe
 1 5 10 15
 Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile
 20 25 30
 Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe
 35 40 45
 Ser Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp
 50 55 60
 Val Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn
 65 70 75 80
 Tyr Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg
 85 90 95
 Thr Arg Val Leu Phe Ile Tyr Xaa
 100

<210> 183
 <211> 198
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (29)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 183
 Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu
 1 5 10 15
 His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln His Xaa Phe Asp Tyr
 20 25 30
 Lys Asp Glu Ser Gly Phe Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr
 35 40 45
 Thr Leu Pro Ser Tyr Asp Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr
 50 55 60
 Ile Pro Leu Val Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp

126

65		70		75		80									
Phe	Asp	Asp	Ala	Asp	Gln	Leu	Arg	Ile	Gly	Asn	Asp	Gly	Ile	Phe	Met
			85						90					95	
Leu	Thr	Phe	Phe	Met	Ala	Phe	Leu	Phe	Asn	Trp	Ile	Gly	Phe	Phe	Leu
			100					105					110		
Ser	Phe	Cys	Leu	Thr	Thr	Ser	Ala	Ala	Gly	Arg	Tyr	Gly	Ala	Ile	Ser
		115				120						125			
Gly	Phe	Gly	Leu	Ser	Leu	Ile	Lys	Trp	Ile	Leu	Ile	Val	Arg	Phe	Ser
		130				135						140			
Thr	Tyr	Phe	Pro	Gly	Tyr	Phe	Asp	Gly	Gln	Tyr	Trp	Leu	Trp	Trp	Val
145				150						155					160
Phe	Leu	Val	Leu	Gly	Phe	Leu	Leu	Phe	Leu	Arg	Gly	Phe	Ile	Asn	Tyr
			165					170						175	
Ala	Lys	Val	Arg	Lys	Met	Pro	Glu	Thr	Phe	Ser	Asn	Leu	Pro	Arg	Thr
			180					185					190		
Arg	Val	Leu	Phe	Ile	Tyr										
			195												

<210> 184
 <211> 70
 <212> PRT
 <213> Homo sapiens

<400> 184
 Met Leu Leu His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln Phe Ser
 1 5 10 15
 Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val
 20 25 30
 Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr
 35 40 45
 Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr
 50 55 60
 Arg Val Leu Phe Ile Tyr
 65 70

<210> 185
 <211> 82
 <212> PRT
 <213> Homo sapiens

<400> 185
 Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe
 1 5 10 15

127

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile
 20 25 30

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe
 35 40 45

Ser Thr Tyr Phe Pro Ala Phe Met Asn Ser Leu Ser Arg Ser Lys Arg
 50 55 60

Thr Pro Ala Gly Ser Glu Ser Arg Cys Arg Thr Gln Arg Asn Asn His
 65 70 75 80

Leu Leu

<210> 186
 <211> 45
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (28)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 186
 Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu
 1 5 10 15

His Leu Thr Ala Ala Phe Leu Gln Arg Ala His Xaa Ile Leu Thr Thr
 20 25 30

Arg Met Ser Leu Gly Phe Gln Ser Pro His Leu Thr Met
 35 40 45

<210> 187
 <211> 34
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (34)
 <223> Xaa equals stop translation

<400> 187
 Met Thr Val Met Asp Pro Lys Gln Met Asn Val Ala Ala Ala Val Trp
 1 5 10 15

Ala Val Val Ser Tyr Val Val Ala Asp Met Glu Glu Met Leu Pro Arg
 20 25 30

Ser Xaa

```
<210> 188
<211> 232
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> SITE  
<222> (232)  
<223> Xaa equals stop translation
```

<400> 188

Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser
1 5 10 15

Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr
20 25 30

Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro
35 40 45

Ile Lys Lys Asn Ser Gly His Ile Tyr Asn Lys Asn Ile Ser Gln Lys
50 55 60

Asp Cys Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly Pro
65 70 75 80

Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg
85 90 95

Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile Leu
100 105 110

Gly Leu Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro Ile
115 120 125

Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp Asp
130 135 140

Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu Ala
145 150 155 160

Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Gly Thr
165 170 175

Ala Ala Leu Glu Ala Ser Ser Pro Arg Ala Ala Lys Ser Leu Ser Leu
180 185 190

Thr Gly Met Leu Ser Ser Ala Asn Trp Gly Ile Glu Phe Lys Val Thr
195 200 205

Arg Lys Lys Gln Ala Asp Asn Trp Lys Gly Thr Asp Trp Val Leu Leu
210 215 220

Gly Phe Ile Leu Ile Pro Cys Xaa
225 230

<210> 189

129

<211> 457

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (457)

<223> Xaa equals stop translation

<400> 189

Met Ala Ala Ala Gly Arg Leu Pro Ser Ser Trp Ala Leu Phe Ser Pro
1 5 10 15

Leu Leu Ala Gly Leu Ala Leu Leu Gly Val Gly Pro Val Pro Ala Arg
20 25 30

Ala Leu His Asn Val Thr Ala Glu Leu Phe Gly Ala Glu Ala Trp Gly
35 40 45

Thr Leu Ala Ala Phe Gly Asp Leu Asn Ser Asp Lys Gln Thr Asp Leu
50 55 60

Phe Val Leu Arg Glu Arg Asn Asp Leu Ile Val Phe Leu Ala Asp Gln
65 70 75 80

Asn Ala Pro Tyr Phe Lys Pro Lys Val Lys Val Ser Phe Lys Asn His
85 90 95

Ser Ala Leu Ile Thr Ser Val Val Pro Gly Asp Tyr Asp Gly Asp Ser
100 105 110

Gln Met Asp Val Leu Leu Thr Tyr Leu Pro Lys Asn Tyr Ala Lys Ser
115 120 125

Glu Leu Gly Ala Val Ile Phe Trp Gly Gln Asn Gln Thr Leu Asp Pro
130 135 140

Asn Asn Met Thr Ile Leu Asn Arg Thr Phe Gln Asp Glu Pro Leu Ile
145 150 155 160

Met Asp Phe Asn Gly Asp Leu Ile Pro Asp Ile Phe Gly Ile Thr Asn
165 170 175

Glu Ser Asn Gln Pro Gln Ile Leu Leu Gly Gly Asn Leu Ser Trp His
180 185 190

Pro Ala Leu Thr Thr Thr Ser Lys Met Arg Ile Pro His Ser His Ala
195 200 205

Phe Ile Asp Leu Thr Glu Asp Phe Thr Ala Asp Leu Phe Leu Thr Thr
210 215 220

Leu Asn Ala Thr Thr Ser Thr Phe Gln Phe Glu Ile Trp Glu Asn Leu
225 230 235 240

Asp Gly Asn Phe Ser Val Ser Thr Ile Leu Glu Lys Pro Gln Asn Met
245 250 255

130

Met Val Val Gly Gln Ser Ala Phe Ala Asp Phe Asp Gly Asp Gly His
 260 265 270

Met Asp His Leu Leu Pro Gly Cys Glu Asp Lys Asn Cys Gln Lys Ser
 275 280 285

Thr Ile Tyr Leu Val Arg Ser Gly Met Lys Gln Trp Val Pro Val Leu
 290 295 300

Gln Asp Phe Ser Asn Lys Gly Thr Leu Trp Gly Phe Val Pro Phe Val
 305 310 315 320

Asp Glu Gln Gln Pro Thr Glu Ile Pro Ile Pro Ile Thr Leu His Ile
 325 330 335

Gly Asp Tyr Asn Met Asp Gly Tyr Pro Asp Ala Leu Val Ile Leu Lys
 340 345 350

Asn Thr Ser Gly Ser Asn Gln Gln Ala Phe Leu Leu Glu Asn Val Pro
 355 360 365

Cys Asn Asn Ala Ser Cys Glu Glu Ala Arg Arg Met Phe Lys Val Tyr
 370 375 380

Trp Glu Leu Thr Asp Leu Asn Gln Ile Lys Asp Ala Met Val Ala Thr
 385 390 395 400

Phe Phe Asp Ile Tyr Glu Asp Gly Ile Leu Asp Ile Val Val Leu Ser
 405 410 415

Lys Gly Tyr Thr Lys Asn Asp Phe Ala Ile His Thr Leu Lys Asn Asn
 420 425 430

Phe Glu Ala Asp Ala Tyr Phe Val Lys Val Ile Val Leu Ser Gly Leu
 435 440 445

Cys Ser Asn Asp Cys Pro Arg Arg Xaa
 450 455

<210> 190

<211> 185

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (185)

<223> Xaa equals stop translation

<400> 190

Met Leu Phe Leu Phe Ser Met Ala Thr Leu Leu Arg Thr Ser Phe Ser
 1 5 10 15

Asp Pro Gly Val Ile Pro Arg Ala Leu Pro Asp Glu Ala Ala Phe Ile
 20 25 30

Glu Met Glu Ile Glu Ala Thr Asn Gly Ala Val Pro Gln Gly Gln Arg

131

35 40 45
 Pro Pro Pro Arg Ile Lys Asn Phe Gln Ile Asn Asn Gln Ile Val Lys
 50 55 60
 Leu Lys Tyr Cys Tyr Thr Cys Lys Ile Phe Arg Pro Pro Arg Ala Ser
 65 70 75 80
 His Cys Ser Ile Cys Asp Asn Cys Val Glu Arg Phe Asp His His Cys
 85 90 95
 Pro Trp Val Gly Asn Cys Val Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr
 100 105 110
 Leu Phe Ile Leu Ser Leu Ser Leu Leu Thr Ile Tyr Val Phe Ala Phe
 115 120 125
 Asn Ile Val Tyr Val Ala Leu Lys Ser Leu Lys Ile Gly Phe Leu Glu
 130 135 140
 Thr Leu Lys Gly Asn Ser Trp Asn Cys Ser Arg Ser Pro His Leu Leu
 145 150 155 160
 Leu Tyr Thr Leu Val Arg Arg Gly Thr Asp Trp Ile Ser Tyr Phe Pro
 165 170 175
 Arg Gly Ser Gln Pro Asp Asn Gln Xaa
 180 185

<210> 191
 <211> 147
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (147)
 <223> Xaa equals stop translation

<400> 191
 Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala
 1 5 10 15
 Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys
 20 25 30
 Phe Arg Gln Ser His Thr Trp Ser Leu Leu Gln Gly Ala Ala His Asp
 35 40 45
 Asp Pro Val Ala Arg Gly Leu Asp Pro Asp Gly Leu Leu Leu Asp
 50 55 60
 Val Val Val Asn Gly Val Val Pro Gly Arg Ala Trp Leu Thr Gln Ile
 65 70 75 80
 Phe Lys Cys Arg Thr Leu Lys Lys His Tyr Val Gln Thr Arg Ala Trp
 85 90 95

132

Pro Ala Val Arg Gly Leu His Thr Ala Leu Leu Pro Gly Arg Pro Pro
 100 105 110

Leu Val Pro Thr Leu Gln Pro Gln His Pro Val Gln Arg Gly Pro Gly
 115 120 125

Pro Pro Ala Pro Ala Gly Ala Ala Pro Ala Gly Leu Ser Tyr Gln Leu
 130 135 140

Gly Leu Xaa
 145

<210> 192

<211> 125

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (125)

<223> Xaa equals stop translation

<400> 192

Met Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Leu Val
 1 5 10 15

Leu Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu
 20 25 30

Ser Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro
 35 40 45

Val Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val
 50 55 60

Met Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu
 65 70 75 80

Pro Ser Cys Thr Asp Gln Cys Phe Arg Ala Lys Pro His Ser Trp Gly
 85 90 95

His Phe Ser Arg Phe Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys
 100 105 110

Thr Arg Asn Glu Phe Arg Leu Pro Ala Arg Ala Asp Xaa
 115 120 125

<210> 193

<211> 52

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (52)

133

<223> Xaa equals stop translation

<400> 193

Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn Cys Ile Phe Phe
 1 5 10 15

Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile Thr Ala Ile Ser
 20 25 30

Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile Phe Phe Pro Cys
 35 40 45

Leu Leu Ala Xaa
 50

<210> 194

<211> 320

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (68)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (115)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (213)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (320)

<223> Xaa equals stop translation

<400> 194

Met Ala Pro Leu Ala Leu His Leu Leu Val Leu Val Pro Ile Leu Leu
 1 5 10 15

Ser Leu Val Ala Ser Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro
 20 25 30

Phe Glu Lys Cys Met Gln Asp Pro Asp Tyr Glu Gln Leu Leu Lys Val
 35 40 45

Thr Ile Leu Glu Ala Asp Asn Arg Ile Gly Gly Arg Ile Phe Thr Tyr
 50 55 60

Arg Asp Gln Xaa Thr Gly Trp Ile Gly Glu Leu Gly Ala Met Arg Met
 65 70 75 80

Pro Ser Ser His Arg Ile Leu His Lys Leu Cys Gln Gly Leu Gly Leu

134
 85 90 95
 Asn Leu Thr Lys Phe Thr Gln Tyr Asp Lys Asn Thr Trp Thr Glu Val
 100 105 110
 His Glu Xaa Lys Leu Arg Asn Tyr Val Val Glu Lys Val Pro Glu Lys
 115 120 125
 Leu Gly Tyr Ala Leu Arg Pro Gln Glu Lys Gly His Ser Pro Glu Asp
 130 135 140
 Ile Tyr Gln Met Ala Leu Asn Gln Ala Leu Lys Asp Leu Lys Ala Leu
 145 150 155 160
 Gly Cys Arg Lys Ala Met Lys Lys Phe Glu Arg His Thr Leu Leu Glu
 165 170 175
 Tyr Leu Leu Gly Glu Gly Asn Leu Ser Arg Pro Ala Val Gln Leu Leu
 180 185 190
 Gly Asp Val Met Ser Glu Asp Gly Phe Phe Tyr Leu Ser Phe Ala Glu
 195 200 205
 Ala Leu Arg Ala Xaa Ser Cys Leu Ser Asp Arg Leu Gln Tyr Ser Arg
 210 215 220
 Ile Val Gly Gly Trp Asp Leu Leu Pro Arg Ala Leu Leu Ser Ser Leu
 225 230 235 240
 Ser Gly Leu Val Leu Leu Asn Ala Pro Val Val Ala Met Thr Gln Gly
 245 250 255
 Pro His Asp Val His Val Gln Ile Glu Thr Ser Pro Pro Ala Arg Asn
 260 265 270
 Leu Lys Val Leu Lys Ala Asp Val Val Leu Leu Thr Ala Ser Gly Pro
 275 280 285
 Ala Val Lys Arg Ile Thr Phe Ser Pro Arg Cys Pro Ala Thr Cys Arg
 290 295 300
 Arg Arg Cys Gly Gly Cys Thr Thr Cys Arg Pro Pro Arg Cys Ser Xaa
 305 310 315 320

<210> 195

<211> 130

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (38)

<223> Xaa equals any of the naturally occurring L-amino acids

135

<220>

<221> SITE

<222> (53)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 195

Pro Phe Cys Ser Gly Phe Phe Pro Ser Leu Trp Ile Tyr Leu Pro Phe
 1 5 10 15

Ile Phe Asn Val Ser Asp Leu Trp Met Gly Ser Leu Ser Gly Cys Ala
 20 25 30

Leu Pro Phe Cys Leu Xaa Val Phe Phe Leu Thr Val Ser Pro Ser Ala
 35 40 45

Val Gly Leu Leu Xaa Phe Ala Gly Gly Pro Leu Gln Thr Leu Phe Ala
 50 55 60

Trp Val Ser Pro Val Glu Ala Ala Glu Gln Gln Arg Leu Leu Pro Val
 65 70 75 80

Leu Ser Ser Gly Ser Phe Val Ser Glu Gly Thr Cys Gln Met Pro Ala
 85 90 95

Arg Ala Leu Leu Tyr Glu Val Ser Val Gly Pro Tyr Trp Glu Ile Pro
 100 105 110

Pro Ser Gln Asp Thr Arg Arg Ser Gly Thr Tyr Leu Arg Arg Gln Ser
 115 120 125

Asp Pro
 130

<210> 196

<211> 108

<212> PRT

<213> Homo sapiens

<400> 196

His Glu Gly Ser Cys Arg Ala Pro Gly Phe Ser Ala His Lys Gly Arg
 1 5 10 15

Gly Cys Pro Ser Pro Arg Met Thr Leu Pro Ser Arg Ala Leu Ala Ser
 20 25 30

Leu Gly Val Gly Val Trp Gly Met Leu Arg Leu Asn Gln Val Thr Val
 35 40 45

Ser Cys Gly Gly Ser Arg Trp Ser Ser Arg Val Ala Leu Gly Ala Phe
 50 55 60

Ser Trp Val Cys Gly Val Ala Leu Val Leu Gln Pro Ser Gly Gly Gly
 65 70 75 80

Leu Gly Leu Thr Ser Pro Ser Glu Gly Cys Trp Glu Gly Glu Leu Ala
 85 90 95

136

Leu Ala Val Leu Arg Ala Pro Gly Gly Ser Pro Ser
 100 105

<210> 197
 <211> 104
 <212> PRT
 <213> Homo sapiens

<400> 197
 Ile Pro Leu Thr Leu Pro Gly Ile Phe Leu Leu Ile Arg Leu Phe Trp
 1 5 10 15
 Arg Leu Gly Gln Ser Ile Cys Gly Pro Gly Lys Leu Val Leu Trp Pro
 20 25 30
 Gln Phe Cys Cys Gly Cys Ala Val Ile Ser Gly His Cys Val Pro Arg
 35 40 45
 Gly Met Pro Ser Ser Trp Leu Pro Gly Cys Phe Val Leu Leu Cys Leu
 50 55 60
 Val Ala Val Gly Cys Gln Leu Arg Glu Trp Gly Val Gly Gly Val Ser
 65 70 75 80
 Ala Val Gly Leu Leu Ala Leu Pro His Leu Gln Val Leu Gly Met Arg
 85 90 95
 Gly Arg Gly Leu Ile Ser Gly Gly
 100

<210> 198
 <211> 237
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (142)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 198
 Gly Pro Ala Gly Lys Glu Ala Trp Ile Trp Ser Trp Leu Leu Pro Ser
 1 5 10 15
 Pro Gly Pro Ala Pro Leu Pro Ser Ala Ser Trp Gly Leu Cys Gly Asp
 20 25 30
 Ala Pro Arg Ala Ala Ala Arg Gly Pro Val Glu Pro Gly Ala Ala Arg
 35 40 45
 Met Ala Leu Leu Ser Arg Pro Ala Leu Thr Leu Leu Leu Leu Met
 50 55 60
 Ala Ala Val Val Arg Cys Gln Glu Gln Ala Gln Thr Thr Asp Trp Arg
 65 70 75 80

137

Ala Thr Leu Lys Thr Ile Arg Asn Gly Val His Lys Ile Asp Thr Tyr
 85 90 95

Leu Asn Ala Ala Leu Asp Leu Leu Gly Gly Glu Asp Gly Leu Cys Gln
 100 105 110

Tyr Lys Cys Ser Asp Gly Ser Lys Pro Phe Pro Arg Tyr Gly Tyr Lys
 115 120 125

Pro Ser Pro Pro Asn Gly Cys Gly Ser Pro Leu Phe Gly Xaa His Leu
 130 135 140

Asn Ile Gly Ile Pro Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg
 145 150 155 160

Cys Tyr Glu Thr Cys Gly Lys Ser Lys Asn Asp Cys Asp Glu Glu Phe
 165 170 175

Gln Tyr Cys Leu Ser Lys Ile Cys Arg Asp Val Gln Lys Thr Leu Gly
 180 185 190

Leu Thr Gln His Val Gln Ala Cys Glu Thr Thr Val Glu Leu Leu Phe
 195 200 205

Asp Ser Val Ile His Leu Gly Cys Lys Pro Tyr Leu Asp Ser Gln Arg.
 210 215 220

Ala Ala Cys Arg Cys His Tyr Glu Glu Lys Thr Asp Leu
 225 230 235

<210> 199
 <211> 8
 <212> PRT
 <213> Homo sapiens

<400> 199
 Cys Cys Asn Gln His Asp Arg Cys
 1 5

<210> 200
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 200
 Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg Cys Tyr Glu Thr
 1 5 10 15

<210> 201
 <211> 16
 <212> PRT
 <213> Homo sapiens

<400> 201
 Leu Thr Lys Cys Cys Asn Gln His Asp Arg Cys Tyr Glu Thr Cys Gly

138

1

5

10

15

<210> 202

<211> 260

<212> PRT

<213> Homo sapiens

<400> 202

Gly Thr Ser Ser Ala Arg Pro Arg Gly Ala Leu Pro Gly Gly Ser Ala
 1 5 10 15

Pro Ser Ala Pro His Gly Gln Leu Pro Gly Arg Ala Gln Pro Ala Pro
 20 25 30

Val Ser Gly Pro Pro Pro Thr Ser Gly Leu Cys His Phe Asp Pro Ala
 35 40 45

Ala Pro Trp Pro Leu Trp Pro Gly Pro Trp Gln Leu Pro Pro His Pro
 50 55 60

Gln Asp Trp Pro Ala His Pro Asp Ile Pro Gln Asp Trp Val Ser Phe
 65 70 75 80

Leu Arg Ser Phe Gly Gln Leu Thr Leu Cys Pro Arg Asn Gly Thr Val
 85 90 95

Thr Gly Lys Trp Arg Gly Ser His Val Val Gly Leu Leu Thr Thr Leu
 100 105 110

Asn Phe Gly Asp Gly Pro Asp Arg Asn Lys Thr Arg Thr Phe Gln Ala
 115 120 125

Thr Val Leu Gly Ser Gln Met Gly Leu Lys Gly Ser Ser Ala Gly Gln
 130 135 140

Leu Val Leu Ile Thr Ala Arg Val Thr Thr Glu Arg Thr Ala Gly Thr
 145 150 155 160

Cys Leu Tyr Phe Ser Ala Val Pro Gly Ile Leu Pro Ser Ser Gln Pro
 165 170 175

Pro Ile Ser Cys Ser Glu Glu Gly Ala Gly Asn Ala Thr Leu Ser Pro
 180 185 190

Arg Met Gly Glu Glu Cys Val Ser Val Trp Ser His Glu Gly Leu Val
 195 200 205

Leu Thr Lys Leu Leu Thr Ser Glu Glu Leu Ala Leu Cys Gly Ser Arg
 210 215 220

Leu Leu Val Leu Gly Ser Phe Leu Leu Leu Phe Cys Gly Leu Leu Cys
 225 230 235 240

Cys Val Thr Ala Met Cys Phe His Pro Arg Arg Glu Ser His Trp Ser

139
245 250 255
Arg Thr Arg Leu
260

<210> 203
<211> 80
<212> PRT
<213> Homo sapiens

<400> 203
Ala Arg Ala Pro Pro Gly Pro Glu Gly Leu Ser Pro Glu Ala Gln Pro
1 5 10 15
Pro Leu Leu Pro Met Gly Asn Cys Gln Ala Gly His Asn Leu His Leu
20 25 30
Cys Leu Ala His His Pro Pro Leu Val Cys Ala Thr Leu Ile Leu Leu
35 40 45
Leu Leu Gly Leu Ser Gly Leu Gly Leu Gly Ser Phe Leu Leu Thr His
50 55 60
Arg Thr Gly Leu Arg Thr Leu Thr Ser Pro Arg Thr Gly Ser Leu Phe
65 70 75 80

<210> 204
<211> 224
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (6)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (9)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (22)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (143)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE

140

<222> (186)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 204

Arg Phe Leu Ser Val Xaa Pro Gln Xaa Glu Val Pro Phe Leu Leu His
1 5 10 15

Pro Cys Val Cys Phe Xaa Gly Gly His Pro Ser Leu Leu Pro Asp Pro
20 25 30

Cys Arg Ala Val Gly Gly Gly Trp Glu Ala Pro Arg Cys Cys Leu His
35 40 45

Glu Ala Leu Cys Gln Ser Leu Gly Cys Lys Ala Glu Glu Ile Val Ser
50 55 60

Val Ser Glu Ser Ser Ser Ala Gln Arg Cys Trp Tyr Leu Leu Arg Gly
65 70 75 80

Arg Lys Ala Gly Gly Arg Gly Pro Ala Ser Pro Val Leu Phe Ala Leu
85 90 95

Met Arg Leu Glu Ser Leu Cys His Leu Cys Leu Ala Cys Leu Phe Phe
100 105 110

Arg Leu Pro Ala Thr Arg Thr Val Tyr Cys Met Asn Glu Ala Glu Ile
115 120 125

Val Asp Val Ala Leu Gly Ile Leu Ile Glu Ser Arg Lys Gln Xaa Lys
130 135 140

Ala Cys Glu Gln Pro Ala Leu Ala Gly Ala Asp Asn Pro Glu His Ser
145 150 155 160

Pro Pro Cys Ser Val Ser Pro His Thr Ser Ser Gly Ser Ser Ser Glu
165 170 175

Glu Glu Asp Ser Gly Lys Gln Ala Leu Xaa Pro Gly Leu Ser Pro Ser
180 185 190

Gln Arg Pro Gly Gly Ser Ser Ser Ala Cys Ser Arg Ser Pro Glu Glu
195 200 205

Glu Glu Glu Glu Asp Val Leu Lys Tyr Val Arg Glu Ile Phe Phe Ser
210 215 220

<210> 205

<211> 199

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (35)

141

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (103)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (191)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 205

Val	Pro	Gly	Trp	Pro	Arg	Ala	Cys	Ser	Pro	Cys	Gln	Ala	Asp	Ser	Pro
1				5					10					15	

Arg	Ala	His	Pro	Pro	Lys	Leu	Arg	Gly	Ile	Leu	Arg	Trp	Ala	Pro	Val
			20					25					30		

Pro	Leu	Xaa	Cys	Ala	Ala	Leu	Cys	Pro	Pro	Leu	Asp	Ser	Gly	Met	Ser
		35					40					45			

Met	Ala	Ala	Cys	Pro	Glu	Ala	Pro	Glu	Pro	Ser	Phe	Leu	Arg	Glu	Val
	50					55						60			

Pro	Ser	Ser	Pro	Ala	Ser	Thr	Gln	Trp	His	Arg	Pro	Cys	Asn	Phe	Arg
65					70					75					80

Gln	Val	Glu	Ala	Asn	Pro	Arg	Lys	Glu	Pro	Lys	Asn	Leu	Val	Trp	Arg
				85					90					95	

Asp	Val	Ser	Leu	Gly	Gln	Xaa	Ser	Arg	Thr	Pro	Arg	Gly	Ser	Gly	Leu
			100					105					110		

Glu	Leu	Val	Arg	Val	Cys	Gly	Gly	Gly	Met	Gln	Arg	Asp	Lys	Thr	Val
		115					120						125		

Val	Glu	Glu	Arg	Val	Gly	Glu	Glu	Arg	Glu	Arg	Glu	Arg	Glu	Arg	Glu
	130					135						140			

Ser	Leu	Gly	Gly	Ala	Gly	Lys	His	Gly	Glu	Met	Arg	Cys	Val	Tyr	Val
145					150					155					160

Arg	Glu	Ser	Val	Gly	Ala	Pro	Gly	Arg	Ala	Gly	Gly	Gly	Gly	Asn	Gly
				165					170					175	

Val	Asn	Ser	Val	Gly	Cys	Val	Arg	Thr	Val	His	Ser	Gly	Ser	Xaa	Pro
			180					185					190		

Pro	Pro	Ser	Ala	Gly	Val	Ser
						195

<210> 206

<211> 174

<212> PRT

<213> Homo sapiens

142

<400> 206

Thr Arg Pro Gly Lys Glu Leu Asn Leu Val Phe Gly Leu Gln Leu Ser
 1 5 10 15
 Met Ala Arg Ile Gly Ser Thr Val Asn Met Asn Leu Met Gly Trp Leu
 20 25 30
 Tyr Ser Lys Ile Glu Ala Leu Leu Gly Ser Ala Gly His Thr Thr Leu
 35 40 45
 Gly Ile Thr Leu Met Ile Gly Gly Ile Thr Cys Ile Leu Ser Leu Ile
 50 55 60
 Cys Ala Leu Ala Leu Ala Tyr Leu Asp Gln Arg Ala Glu Arg Ile Leu
 65 70 75 80
 His Lys Glu Gln Gly Lys Thr Gly Glu Val Ile Lys Leu Thr Asp Val
 85 90 95
 Lys Asp Phe Ser Leu Pro Leu Trp Leu Ile Phe Ile Ile Cys Val Cys
 100 105 110
 Tyr Tyr Val Ala Val Phe Pro Phe Ile Gly Leu Gly Lys Val Phe Phe
 115 120 125
 Thr Glu Lys Phe Gly Phe Ser Ser Gln Ala Ala Ser Ala Ile Asn Ser
 130 135 140
 Val Val Tyr Val Ile Ser Ala Pro Met Ser Pro Val Phe Gly Leu Leu
 145 150 155 160
 Val Asp Lys Thr Gly Lys Asn Ile Ile Trp Val Leu Cys Ala
 165 170

<210> 207

<211> 31

<212> PRT

<213> Homo sapiens

<400> 207

Cys Lys Asp Leu Cys Ser Arg Val Tyr Leu Leu Thr Leu Ser Pro Leu
 1 5 10 15
 Leu Ser Tyr Asp Pro Ala Thr Ser His Ser Pro Arg Asn Thr Gln
 20 25 30

<210> 208

<211> 369

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (78)

<223> Xaa equals any of the naturally occurring L-amino acids

143

<400> 208

Ile Ile Cys Glu Cys Trp Glu Glu Glu Cys Gln Ser Cys Arg Leu Lys
 1 5 10 15
 Ile Thr Gln Pro Arg Glu Ile Cys Arg Met Asp Phe Leu Val Leu Phe
 20 25 30
 Leu Phe Tyr Leu Ala Ser Val Leu Met Gly Leu Val Leu Ile Cys Val
 35 40 45
 Cys Ser Lys Thr His Ser Leu Lys Gly Leu Ala Arg Gly Gly Ala Gln
 50 55 60
 Ile Phe Ser Cys Ile Ile Pro Glu Cys Leu Gln Arg Ala Xaa His Gly
 65 70 75 80
 Leu Leu His Tyr Leu Phe His Thr Arg Asn His Thr Phe Ile Val Leu
 85 90 95
 His Leu Val Leu Gln Gly Met Val Tyr Thr Glu Tyr Thr Trp Glu Val
 100 105 110
 Phe Gly Tyr Cys Gln Glu Leu Glu Leu Ser Leu His Tyr Leu Leu Leu
 115 120 125
 Pro Tyr Leu Leu Leu Gly Val Asn Leu Phe Phe Phe Thr Leu Thr Cys
 130 135 140
 Gly Thr Asn Pro Gly Ile Ile Thr Lys Ala Asn Glu Leu Leu Phe Leu
 145 150 155 160
 His Val Tyr Glu Phe Asp Glu Val Met Phe Pro Lys Asn Val Arg Cys
 165 170 175
 Ser Thr Cys Asp Leu Arg Lys Pro Ala Arg Ser Lys His Cys Ser Val
 180 185 190
 Cys Asn Trp Cys Val His Arg Phe Asp His His Cys Val Trp Val Asn
 195 200 205
 Asn Cys Ile Gly Ala Trp Asn Ile Arg Tyr Phe Leu Ile Tyr Val Leu
 210 215 220
 Thr Leu Thr Ala Ser Ala Ala Thr Val Ala Ile Val Ser Thr Thr Phe
 225 230 235 240
 Leu Val His Leu Val Val Met Ser Asp Leu Tyr Gln Glu Thr Tyr Ile
 245 250 255
 Asp Asp Leu Gly His Leu His Val Met Asp Thr Val Phe Leu Ile Gln
 260 265 270
 Tyr Leu Phe Leu Thr Phe Pro Arg Ile Val Phe Met Leu Gly Phe Val
 275 280 285
 Val Val Leu Ser Phe Leu Leu Gly Gly Tyr Leu Leu Phe Val Leu Tyr
 290 295 300

144

Leu Ala Ala Thr Asn Gln Thr Thr Asn Glu Trp Tyr Arg Gly Asp Trp
 305 310 315 320

Ala Trp Cys Gln Arg Cys Pro Leu Val Ala Trp Pro Pro Ser Ala Glu
 325 330 335

Pro Gln Val His Arg Asn Ile His Ser His Gly Leu Arg Ser Asn Leu
 340 345 350

Gln Glu Ile Phe Leu Pro Ala Phe Pro Cys His Glu Arg Lys Lys Gln
 355 360 365

Glu

<210> 209

<211> 147

<212> PRT

<213> Homo sapiens

<400> 209

Leu Leu Ser Phe Lys Ile Arg Gly Leu Arg Thr Glu Asp Ala Gly Trp
 1 5 10 15

Ala Gln Ser Ser Ser Gly Gly Leu Cys Val Arg Gly Asp Ala Phe Trp
 20 25 30

Met Pro Ser Ser Ser Ser Gly Leu Gly Ser Pro Ser Arg Pro Pro Ser
 35 40 45

Ser Phe Leu Cys Leu Leu Leu Leu Leu Leu Pro Pro Ala Ala Leu Ala
 50 55 60

Leu Leu Leu Phe Phe Leu Asp Phe Phe Pro Pro Arg Ala Ala Val Ser
 65 70 75 80

Pro Phe Leu Pro Asp His Cys Ser Ala Arg Gln Pro Arg Val Trp Arg
 85 90 95

Arg Glu Thr Leu Asn Arg Ser Ala Ser Gly Leu Gly Cys Trp Ala Arg
 100 105 110

Ser Thr Glu Gln Gly Ala Val Gly Val Ala Thr Gly Thr Val Leu Asp
 115 120 125

Ile Ser Leu Pro Ala Ser Cys Leu Ser Leu Trp Pro Pro Gly Pro Ser
 130 135 140

Gly Gly Ile
 145

<210> 210

<211> 143

<212> PRT

<213> Homo sapiens

145

<400> 210

Gln Leu Gly Leu Cys Leu Thr Ser Ala Ser Leu Pro Pro Ala Ser Arg
 1 5 10 15

Cys Gly His Gln Ala Pro Leu Gly Ala Ser Asp Leu Ser Ala His His
 20 25 30

Ser Ala Pro Gly Phe Ser Asp Ser Tyr Phe Thr Met Ser Cys Gln Ser
 35 40 45

Ser Leu Ser Arg Ala Glu Ile Leu Gln Cys Pro Leu Val Pro Ser Val
 50 55 60

Ser Pro Pro Thr His Leu Pro Gln Gly Arg Ala Asn Lys Ser Ser Arg
 65 70 75 80

Ala Ser Leu Pro Leu Leu Pro Gln Thr His Trp Cys Leu Phe Pro Ser
 85 90 95

Ala Arg Gly Trp Arg Arg Gly Ile Gln Ser Gly Leu Pro Pro Gly Gly
 100 105 110

Ser Cys Thr Ser Pro Arg Ser Pro Pro Gln Thr Leu His Gln His Ile
 115 120 125

Thr Leu Val Asn His Asn Thr Ser Tyr Trp Gln Ser Pro Ser Thr
 130 135 140

<210> 211

<211> 160

<212> PRT

<213> Homo sapiens

<400> 211

His Gln Pro Pro Cys Leu Leu Pro Leu Ala Val Ala Thr Arg Pro Leu
 1 5 10 15

Trp Gly His Leu Thr Cys Leu Pro Ile Ile Leu His Leu Val Ser Val
 20 25 30

Thr Leu Thr Ser Pro Cys Leu Ala Asn Gln Ala Phe Gln Gly Gln Arg
 35 40 45

Ser Tyr Asn Ala Leu Trp Cys Pro Leu Phe Leu Leu Leu Pro Thr Ser
 50 55 60

Pro Lys Gly Glu Gln Thr Asn His Pro Glu Pro Ala Cys Pro Cys Phe
 65 70 75 80

Pro Lys Leu Thr Gly Val Phe Ser Leu Gln His Val Val Gly Ala Glu
 85 90 95

Glu Phe Ser Gln Val Phe Leu Leu Val Asp Pro Val Pro Val Leu Asp
 100 105 110

His Leu Leu Lys Leu Phe Thr Ser Thr Ser His Leu Leu Ile Ile Ile
 115 120 125

146

Pro His Ile Gly Lys Ala Pro Ala Pro Asp Ser Leu Leu Glu Glu Leu
 130 135 140

Ser Leu Ser Leu Ala Thr His Cys Lys Val Ala Val Ala Arg Phe Thr
 145 150 155 160

<210> 212
 <211> 157
 <212> PRT
 <213> Homo sapiens

<400> 212
 Met Ala Ala Glu Gly Ser Arg Phe Ser Ser Gln Ser Pro Gly Leu Val
 1 5 10 15

Asp Arg Gln Gly Pro Lys Cys Asp Pro Ser Arg Leu Val Ser Pro Trp
 20 25 30

Gly Arg His Gly Leu Arg Ile Leu Gln Ile Gly His His His Gly Arg
 35 40 45

Asp Gly Gln His Glu Ala Thr His His Leu Leu Arg Val Leu Arg Ala
 50 55 60

Pro Arg Val Gly Lys Ala Asp Glu Gly Ala Val Asp Ser Asp Pro Ser
 65 70 75 80

Thr Pro Leu Gln Leu Lys His Glu Ala Ala His Ala Glu Asp His Ala
 85 90 95

Gln Gln Val His Val Val Arg Arg Arg Val Val Gln Gly Arg Val Thr
 100 105 110

Phe Ala Arg Arg Gly Leu Val Pro Gln His Phe Val Arg Pro Pro Trp
 115 120 125

Val Arg His Ile Val Ser Gly His Ser Glu Ser Lys Ala Arg Ser Arg
 130 135 140

Leu Phe Arg Cys Arg Asn Arg Ser Phe Arg Arg Ala Ser
 145 150 155

<210> 213
 <211> 38
 <212> PRT
 <213> Homo sapiens

<400> 213
 Arg Leu Val Ser Pro Trp Gly Arg His Gly Leu Arg Ile Leu Gln Ile
 1 5 10 15

Gly His His His Gly Arg Asp Gly Gln His Glu Ala Thr His His Leu

147

20

25

30

Leu Arg Val Leu Arg Ala
35

<210> 214
<211> 12
<212> PRT
<213> Homo sapiens

<400> 214
Pro Thr Asp Val Leu Lys Ile Arg Met Gln Ala Gln
1 5 10

<210> 215
<211> 7
<212> PRT
<213> Homo sapiens

<400> 215
Thr Tyr Glu Gln Leu Lys Arg
1 5

<210> 216
<211> 137
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (22)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (33)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (71)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 216
Arg Pro Arg Pro Ser Ala Ser Ser Leu Ala Arg Ser Ala Ser Leu Leu
1 5 10 15

Pro Ala Ala His Gly Xaa Gly Val Gly Gly Ala Gly Gly Gly Ser Ser
20 25 30

Xaa Leu Arg Ser Arg Tyr Gln Gln Leu Gln Asn Glu Glu Glu Ser Gly
35 40 45

Glu Pro Glu Gln Ala Ala Gly Asp Ala Pro Pro Tyr Ser Ser Ile
50 55 60

148

Ser Ala Glu Ser Ala His Xaa Phe Asp Tyr Lys Asp Glu Ser Gly Phe
65 70 75 80

Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr Thr Leu Pro Ser Tyr Asp
85 90 95

Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr Ile Pro Leu Val Pro Gly
100 105 110

Arg Asp Glu Asp Phe Val Gly Arg Asp Asp Phe Asp Asp Ala Asp Gln
115 120 125

Leu Arg Ile Gly Asn Asp Gly Ile Phe
130 135

<210> 217

<211> 20

<212> PRT

<213> Homo sapiens

<400> 217

Arg Tyr Gln Gln Leu Gln Asn Glu Glu Glu Ser Gly Glu Pro Glu Gln
1 5 10 15

Ala Ala Gly Asp
20

<210> 218

<211> 22

<212> PRT

<213> Homo sapiens

<400> 218

Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp Phe Asp Asp Ala
1 5 10 15

Asp Gln Leu Arg Ile Gly
20

<210> 219

<211> 103

<212> PRT

<213> Homo sapiens

<400> 219

Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe
1 5 10 15

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile
20 25 30

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe
35 40 45

149

Ser Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp
 50 55 60
 Val Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn
 65 70 75 80
 Tyr Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg
 85 90 95
 Thr Arg Val Leu Phe Ile Tyr
 100

<210> 220
 <211> 198
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (29)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 220
 Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu
 1 5 10 15
 His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln His Xaa Phe Asp Tyr
 20 25 30
 Lys Asp Glu Ser Gly Phe Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr
 35 40 45
 Thr Leu Pro Ser Tyr Asp Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr
 50 55 60
 Ile Pro Leu Val Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp
 65 70 75 80
 Phe Asp Asp Ala Asp Gln Leu Arg Ile Gly Asn Asp Gly Ile Phe Met
 85 90 95
 Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe Leu
 100 105 110
 Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile Ser
 115 120 125
 Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe Ser
 130 135 140
 Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val
 145 150 155 160
 Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr
 165 170 175
 Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr

150

180

185

190

Arg Val Leu Phe Ile Tyr
195

<210> 221
<211> 70
<212> PRT
<213> Homo sapiens

<400> 221
Met Leu Leu His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln Phe Ser
1 5 10 15
Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val
20 25 30
Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr
35 40 45
Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr
50 55 60
Arg Val Leu Phe Ile Tyr
65 70

<210> 222
<211> 82
<212> PRT
<213> Homo sapiens

<400> 222
Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe
1 5 10 15
Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile
20 25 30
Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe
35 40 45
Ser Thr Tyr Phe Pro Ala Phe Met Asn Ser Leu Ser Arg Ser Lys Arg
50 55 60
Thr Pro Ala Gly Ser Glu Ser Arg Cys Arg Thr Gln Arg Asn Asn His
65 70 75 80
Leu Leu

<210> 223
<211> 45
<212> PRT
<213> Homo sapiens

151

<220>

<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 223

Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu
 1 5 10 15

His Leu Thr Ala Ala Phe Leu Gln Arg Ala His Xaa Ile Leu Thr Thr
 20 25 30

Arg Met Ser Leu Gly Phe Gln Ser Pro His Leu Thr Met
 35 40 45

<210> 224

<211> 33

<212> PRT

<213> Homo sapiens

<400> 224

Met Thr Val Met Asp Pro Lys Gln Met Asn Val Ala Ala Ala Val Trp
 1 5 10 15

Ala Val Val Ser Tyr Val Val Ala Asp Met Glu Glu Met Leu Pro Arg
 20 25 30

Ser

<210> 225

<211> 189

<212> PRT

<213> Homo sapiens

<400> 225

Pro Arg Val Arg Ser Arg Glu Pro Val Ala Gly Ala Pro Gly Cys Gly
 1 5 10 15

Thr Ala Gly Pro Pro Ala Met Ala Thr Leu Trp Gly Gly Leu Leu Arg
 20 25 30

Leu Gly Ser Leu Leu Ser Leu Ser Cys Leu Ala Leu Ser Val Leu Leu
 35 40 45

Leu Ala His Cys Gln Thr Pro Pro Ser Asp Cys Leu His Val Val Glu
 50 55 60

Pro Met Pro Val Arg Gly Pro Asp Val Glu Ala Tyr Cys Leu Arg Cys
 65 70 75 80

Glu Cys Lys Tyr Glu Glu Arg Ser Ser Val Thr Ile Lys Val Thr Ile
 85 90 95

Ile Ile Tyr Leu Ser Ile Leu Gly Leu Leu Leu Tyr Met Val Tyr
 100 105 110

152

Leu Thr Leu Val Glu Pro Ile Leu Lys Arg Arg Leu Phe Gly His Ala
115 120 125

Gln Leu Ile Gln Ser Asp Asp Asp Ile Gly Asp His Gln Pro Phe Ala
130 135 140

Asn Ala His Asp Val Leu Ala Arg Ser Arg Ser Arg Ala Asn Val Leu
145 150 155 160

Asn Lys Val Glu Tyr Ala Gln Gln Arg Trp Lys Leu Gln Val Gln Glu
165 170 175

Gln Arg Lys Ser Val Phe Asp Arg His Val Val Leu Ser
180 185

<210> 226

<211> 231

<212> PRT

<213> Homo sapiens

<400> 226

Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser
1 5 10 15

Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr
20 25 30

Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro
35 40 45

Ile Lys Lys Asn Ser Gly His Ile Tyr Asn Lys Asn Ile Ser Gln Lys
50 55 60

Asp Cys Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly Pro
65 70 75 80

Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg
85 90 95

Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile Leu
100 105 110

Gly Leu Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro Ile
115 120 125

Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp Asp
130 135 140

Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu Ala
145 150 155 160

Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Gly Thr
165 170 175

Ala Ala Leu Glu Ala Ser Ser Pro Arg Ala Ala Lys Ser Leu Ser Leu
180 185 190

153

Thr Gly Met Leu Ser Ser Ala Asn Trp Gly Ile Glu Phe Lys Val Thr
 195 200 205
 Arg Lys Lys Gln Ala Asp Asn Trp Lys Gly Thr Asp Trp Val Leu Leu
 210 215 220
 Gly Phe Ile Leu Ile Pro Cys
 225 230
 <210> 227
 <211> 456
 <212> PRT
 <213> Homo sapiens
 <400> 227
 Met Ala Ala Ala Gly Arg Leu Pro Ser Ser Trp Ala Leu Phe Ser Pro
 1 5 10 15
 Leu Leu Ala Gly Leu Ala Leu Leu Gly Val Gly Pro Val Pro Ala Arg
 20 25 30
 Ala Leu His Asn Val Thr Ala Glu Leu Phe Gly Ala Glu Ala Trp Gly
 35 40 45
 Thr Leu Ala Ala Phe Gly Asp Leu Asn Ser Asp Lys Gln Thr Asp Leu
 50 55 60
 Phe Val Leu Arg Glu Arg Asn Asp Leu Ile Val Phe Leu Ala Asp Gln
 65 70 75 80
 Asn Ala Pro Tyr Phe Lys Pro Lys Val Lys Val Ser Phe Lys Asn His
 85 90 95
 Ser Ala Leu Ile Thr Ser Val Val Pro Gly Asp Tyr Asp Gly Asp Ser
 100 105 110
 Gln Met Asp Val Leu Leu Thr Tyr Leu Pro Lys Asn Tyr Ala Lys Ser
 115 120 125
 Glu Leu Gly Ala Val Ile Phe Trp Gly Gln Asn Gln Thr Leu Asp Pro
 130 135 140
 Asn Asn Met Thr Ile Leu Asn Arg Thr Phe Gln Asp Glu Pro Leu Ile
 145 150 155 160
 Met Asp Phe Asn Gly Asp Leu Ile Pro Asp Ile Phe Gly Ile Thr Asn
 165 170 175
 Glu Ser Asn Gln Pro Gln Ile Leu Leu Gly Gly Asn Leu Ser Trp His
 180 185 190
 Pro Ala Leu Thr Thr Thr Ser Lys Met Arg Ile Pro His Ser His Ala
 195 200 205
 Phe Ile Asp Leu Thr Glu Asp Phe Thr Ala Asp Leu Phe Leu Thr Thr
 210 215 220

154

Leu Asn Ala Thr Thr Ser Thr Phe Gln Phe Glu Ile Trp Glu Asn Leu
 225 230 235 240
 Asp Gly Asn Phe Ser Val Ser Thr Ile Leu Glu Lys Pro Gln Asn Met
 245 250 255
 Met Val Val Gly Gln Ser Ala Phe Ala Asp Phe Asp Gly Asp Gly His
 260 265 270
 Met Asp His Leu Leu Pro Gly Cys Glu Asp Lys Asn Cys Gln Lys Ser
 275 280 285
 Thr Ile Tyr Leu Val Arg Ser Gly Met Lys Gln Trp Val Pro Val Leu
 290 295 300
 Gln Asp Phe Ser Asn Lys Gly Thr Leu Trp Gly Phe Val Pro Phe Val
 305 310 315 320
 Asp Glu Gln Gln Pro Thr Glu Ile Pro Ile Pro Ile Thr Leu His Ile
 325 330 335
 Gly Asp Tyr Asn Met Asp Gly Tyr Pro Asp Ala Leu Val Ile Leu Lys
 340 345 350
 Asn Thr Ser Gly Ser Asn Gln Gln Ala Phe Leu Leu Glu Asn Val Pro
 355 360 365
 Cys Asn Asn Ala Ser Cys Glu Glu Ala Arg Arg Met Phe Lys Val Tyr
 370 375 380
 Trp Glu Leu Thr Asp Leu Asn Gln Ile Lys Asp Ala Met Val Ala Thr
 385 390 395 400
 Phe Phe Asp Ile Tyr Glu Asp Gly Ile Leu Asp Ile Val Val Leu Ser
 405 410 415
 Lys Gly Tyr Thr Lys Asn Asp Phe Ala Ile His Thr Leu Lys Asn Asn
 420 425 430
 Phe Glu Ala Asp Ala Tyr Phe Val Lys Val Ile Val Leu Ser Gly Leu
 435 440 445
 Cys Ser Asn Asp Cys Pro Arg Arg
 450 455

<210> 228

<211> 282

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (144)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

155

<221> SITE

<222> (168)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 228

```

Met Thr Lys Arg Glu Asp Gly Gly Tyr Thr Phe Thr Ala Thr Pro Glu
  1             5             10             15

Asp Phe Pro Lys Lys His Lys Ala Pro Val Ile Asp Ile Gly Ile Ala
          20             25             30

Asn Thr Gly Lys Phe Ile Met Thr Ala Ser Ser Asp Thr Thr Val Leu
      35             40             45

Ile Trp Ser Leu Lys Gly Gln Val Leu Ser Thr Ile Asn Thr Asn Gln
      50             55             60

Met Asn Asn Thr His Ala Ala Val Ser Pro Cys Gly Arg Phe Val Ala
      65             70             75             80

Ser Cys Gly Phe Thr Pro Asp Val Lys Val Trp Glu Val Cys Phe Gly
          85             90             95

Lys Lys Gly Glu Phe Gln Glu Val Val Arg Ala Phe Glu Leu Lys Gly
          100             105             110

His Ser Ala Ala Val His Ser Phe Ala Phe Ser Asn Asp Ser Arg Arg
          115             120             125

Met Ala Ser Val Ser Lys Asp Gly Thr Trp Lys Leu Trp Asp Thr Xaa
          130             135             140

Val Glu Tyr Lys Lys Lys Gln Asp Pro Tyr Leu Leu Lys Thr Gly Arg
          145             150             155             160

Phe Glu Glu Ala Ala Gly Ala Xaa Pro Cys Arg Leu Ala Leu Ser Pro
          165             170             175

Asn Ala Gln Val Leu Ala Leu Ala Ser Gly Ser Ser Ile His Leu Tyr
          180             185             190

Asn Thr Arg Arg Gly Glu Lys Glu Glu Cys Phe Glu Arg Val His Gly
          195             200             205

Glu Cys Ile Ala Asn Leu Ser Phe Asp Ile Thr Gly Arg Phe Leu Ala
          210             215             220

Ser Cys Gly Asp Arg Ala Val Arg Leu Phe His Asn Thr Pro Gly His
          225             230             235             240

Arg Ala Met Val Glu Glu Met Gln Gly His Leu Lys Arg Ala Ser Asn
          245             250             255

Glu Ser Thr Arg Gln Arg Leu Gln Gln Gln Leu Thr Gln Ala Gln Glu
          260             265             270

Thr Leu Lys Ser Leu Gly Ala Leu Lys Lys
          275             280

```


156

<210> 229
 <211> 456
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (17)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (37)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (318)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (342)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 229
 Val Ile Arg His Glu Gly Ser Thr Asn Met Glu Leu Ser Gln Met Ser
 1 5 10 15

Xaa Leu Met Gly Leu Ser Val Leu Leu Gly Leu Leu Ala Leu Met Ala
 20 25 30

Thr Ala Ala Val Xaa Arg Gly Trp Leu Arg Ala Gly Glu Glu Arg Ser
 35 40 45

Gly Arg Pro Ala Cys Gln Lys Ala Asn Gly Phe Pro Pro Asp Lys Ser
 50 55 60

Ser Gly Ser Lys Lys Gln Lys Gln Tyr Gln Arg Ile Arg Lys Glu Lys
 65 70 75 80

Pro Gln Gln His Asn Phe Thr His Arg Leu Leu Ala Ala Ala Leu Lys
 85 90 95

Ser His Ser Gly Asn Ile Ser Cys Met Asp Phe Ser Ser Asn Gly Lys
 100 105 110

Tyr Leu Ala Thr Cys Ala Asp Asp Arg Thr Ile Arg Ile Trp Ser Thr
 115 120 125

Lys Asp Phe Leu Gln Arg Glu His Arg Ser Met Arg Ala Asn Val Glu
 130 135 140

Leu Asp His Ala Thr Leu Val Arg Phe Ser Pro Asp Cys Arg Ala Phe
 145 150 155 160

157

Ile Val Trp Leu Ala Asn Gly Asp Thr Leu Arg Val Phe Lys Met Thr
 165 170 175
 Lys Arg Glu Asp Gly Gly Tyr Thr Phe Thr Ala Thr Pro Glu Asp Phe
 180 185 190
 Pro Lys Lys His Lys Ala Pro Val Ile Asp Ile Gly Ile Ala Asn Thr
 195 200 205
 Gly Lys Phe Ile Met Thr Ala Ser Ser Asp Thr Thr Val Leu Ile Trp
 210 215 220
 Ser Leu Lys Gly Gln Val Leu Ser Thr Ile Asn Thr Asn Gln Met Asn
 225 230 235 240
 Asn Thr His Ala Ala Val Ser Pro Cys Gly Arg Phe Val Ala Ser Cys
 245 250 255
 Gly Phe Thr Pro Asp Val Lys Val Trp Glu Val Cys Phe Gly Lys Lys
 260 265 270
 Gly Glu Phe Gln Glu Val Val Arg Ala Phe Glu Leu Lys Gly His Ser
 275 280 285
 Ala Ala Val His Ser Phe Ala Phe Ser Asn Asp Ser Arg Arg Met Ala
 290 295 300
 Ser Val Ser Lys Asp Gly Thr Trp Lys Leu Trp Asp Thr Xaa Val Glu
 305 310 315 320
 Tyr Lys Lys Lys Gln Asp Pro Tyr Leu Leu Lys Thr Gly Arg Phe Glu
 325 330 335
 Glu Ala Ala Gly Ala Xaa Pro Cys Arg Leu Ala Leu Ser Pro Asn Ala
 340 345 350
 Gln Val Leu Ala Leu Ala Ser Gly Ser Ser Ile His Leu Tyr Asn Thr
 355 360 365
 Arg Arg Gly Glu Lys Glu Glu Cys Phe Glu Arg Val His Gly Glu Cys
 370 375 380
 Ile Ala Asn Leu Ser Phe Asp Ile Thr Gly Arg Phe Leu Ala Ser Cys
 385 390 395 400
 Gly Asp Arg Ala Val Arg Leu Phe His Asn Thr Pro Gly His Arg Ala
 405 410 415
 Met Val Glu Glu Met Gln Gly His Leu Lys Arg Ala Ser Asn Glu Ser
 420 425 430
 Thr Arg Gln Arg Leu Gln Gln Gln Leu Thr Gln Ala Gln Glu Thr Leu
 435 440 445
 Lys Ser Leu Gly Ala Leu Lys Lys
 450 455

158

<210> 230

<211> 363

<212> PRT

<213> Homo sapiens

<400> 230

```

Met Ser Val Met Val Val Arg Lys Lys Val Thr Arg Lys Trp Glu Lys
 1              5              10              15

Leu Pro Gly Arg Asn Thr Phe Cys Cys Asp Gly Arg Val Met Met Ala
              20              25              30

Arg Gln Lys Gly Ile Phe Tyr Leu Thr Leu Phe Leu Ile Leu Gly Thr
              35              40              45

Cys Thr Leu Phe Phe Ala Phe Glu Cys Arg Tyr Leu Ala Val Gln Leu
 50              55              60

Ser Pro Ala Ile Pro Val Phe Ala Ala Met Leu Phe Leu Phe Ser Met
 65              70              75              80

Ala Thr Leu Leu Arg Thr Ser Phe Ser Asp Pro Gly Val Ile Pro Arg
              85              90              95

Ala Leu Pro Asp Glu Ala Ala Phe Ile Glu Met Glu Ile Glu Ala Thr
              100              105              110

Asn Gly Ala Val Pro Gln Gly Gln Arg Pro Pro Pro Arg Ile Lys Asn
              115              120              125

Phe Gln Ile Asn Asn Gln Ile Val Lys Leu Lys Tyr Cys Tyr Thr Cys
              130              135              140

Lys Ile Phe Arg Pro Pro Arg Ala Ser His Cys Ser Ile Cys Asp Asn
              145              150              155              160

Cys Val Glu Arg Phe Asp His His Cys Pro Trp Val Gly Asn Cys Val
              165              170              175

Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr Leu Phe Ile Leu Ser Leu Ser
              180              185              190

Leu Leu Thr Ile Tyr Val Phe Ala Phe Asn Ile Val Tyr Val Ala Leu
              195              200              205

Lys Ser Leu Lys Ile Gly Phe Leu Glu Thr Leu Lys Glu Thr Pro Gly
              210              215              220

Thr Val Leu Glu Val Leu Ile Cys Phe Phe Thr Leu Trp Ser Val Val
              225              230              235              240

Gly Leu Thr Gly Phe His Thr Phe Leu Val Ala Leu Asn Gln Thr Thr
              245              250              255

Asn Glu Asp Ile Lys Gly Ser Trp Thr Gly Lys Asn Arg Val Gln Asn
              260              265              270

Pro Tyr Ser His Gly Asn Ile Val Lys Asn Cys Cys Glu Val Leu Cys

```

159

275 280 285
 Gly Pro Leu Pro Pro Ser Val Leu Asp Arg Arg Gly Ile Leu Pro Leu
 290 295 300
 Glu Glu Ser Gly Ser Arg Pro Pro Ser Thr Gln Glu Thr Ser Ser Ser
 305 310 315 320
 Leu Leu Pro Gln Ser Pro Ala Pro Thr Glu Leu Asn Ser Asn Glu Met
 325 330 335
 Pro Glu Asp Ser Ser Thr Pro Glu Glu Met Pro Pro Pro Glu Pro Pro
 340 345 350
 Glu Pro Pro Gln Glu Ala Ala Glu Ala Glu Lys
 355 360

 <210> 231
 <211> 184
 <212> PRT
 <213> Homo sapiens

 <400> 231
 Met Leu Phe Leu Phe Ser Met Ala Thr Leu Leu Arg Thr Ser Phe Ser
 1 5 10 15
 Asp Pro Gly Val Ile Pro Arg Ala Leu Pro Asp Glu Ala Ala Phe Ile
 20 25 30
 Glu Met Glu Ile Glu Ala Thr Asn Gly Ala Val Pro Gln Gly Gln Arg
 35 40 45
 Pro Pro Pro Arg Ile Lys Asn Phe Gln Ile Asn Asn Gln Ile Val Lys
 50 55 60
 Leu Lys Tyr Cys Tyr Thr Cys Lys Ile Phe Arg Pro Pro Arg Ala Ser
 65 70 75 80
 His Cys Ser Ile Cys Asp Asn Cys Val Glu Arg Phe Asp His His Cys
 85 90 95
 Pro Trp Val Gly Asn Cys Val Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr
 100 105 110
 Leu Phe Ile Leu Ser Leu Ser Leu Leu Thr Ile Tyr Val Phe Ala Phe
 115 120 125
 Asn Ile Val Tyr Val Ala Leu Lys Ser Leu Lys Ile Gly Phe Leu Glu
 130 135 140
 Thr Leu Lys Gly Asn Ser Trp Asn Cys Ser Arg Ser Pro His Leu Leu
 145 150 155 160
 Leu Tyr Thr Leu Val Arg Arg Gly Thr Asp Trp Ile Ser Tyr Phe Pro
 165 170 175
 Arg Gly Ser Gln Pro Asp Asn Gln

160

180

<210> 232
 <211> 52
 <212> PRT
 <213> Homo sapiens

<400> 232
 Leu His Glu Cys Leu Pro Gly Ser Ile Ser Tyr Leu His Pro Arg Thr
 1 5 10 15
 Pro Trp Leu Cys Leu Pro Pro Gln His Leu Ser Phe Ser Thr Phe Ser
 20 25 30
 Pro Pro Trp Gln Pro Ala Met Ser Pro Val Pro Gly Thr Gly Gly Pro
 35 40 45
 Pro Cys Gly Leu
 50

<210> 233
 <211> 177
 <212> PRT
 <213> Homo sapiens

<400> 233
 Met Leu Pro Leu Leu Ile Ile Cys Leu Leu Pro Ala Ile Glu Gly Lys
 1 5 10 15
 Asn Cys Leu Arg Cys Trp Pro Glu Leu Ser Ala Leu Ile Asp Tyr Asp
 20 25 30
 Leu Gln Ile Leu Trp Val Thr Pro Gly Pro Pro Thr Glu Leu Ser Gln
 35 40 45
 Ser Ile His Ser Leu Phe Leu Glu Asp Asn Asn Phe Leu Lys Pro Trp
 50 55 60
 Tyr Leu Asp Arg Asp His Leu Glu Glu Glu Thr Ala Lys Phe Phe Thr
 65 70 75 80
 Gln Val His Gln Ala Ile Lys Thr Leu Arg Asp Asp Lys Thr Val Leu
 85 90 95
 Leu Glu Glu Ile Tyr Thr His Lys Asn Leu Phe Thr Glu Arg Leu Asn
 100 105 110
 Lys Ile Ser Asp Gly Leu Lys Glu Lys Gly Ala Pro Pro Leu His Glu
 115 120 125
 Cys Leu Pro Gly Ser Ile Ser Tyr Leu His Pro Arg Thr Pro Trp Leu
 130 135 140
 Cys Leu Pro Pro Gln His Leu Ser Phe Ser Thr Phe Ser Pro Pro Trp
 145 150 155 160

161

Gln Pro Ala Met Ser Pro Val Pro Gly Thr Gly Gly Pro Pro Cys Gly
 165 170 175

Leu

<210> 234

<211> 95

<212> PRT

<213> Homo sapiens

<400> 234

Pro Pro Val Pro Pro Trp Ile Ser Leu Pro Leu Thr Gly Ser Pro Pro
 1 5 10 15

Arg Pro Gly Phe Val Pro Val Ser Pro Phe Cys Phe Ser Pro Met Thr
 20 25 30

Asn Gly His Gln Val Leu Leu Leu Leu Leu Thr Ser Ala Val Ala
 35 40 45

Ala Gly Pro Trp Pro Gln Val His Ala Gly Gln Trp Gly Trp Met Cys
 50 55 60

Leu Pro Pro Gly Leu Pro Ser Val Gln Ala Arg Ser Gly Leu Gly Gly
 65 70 75 80

Leu Pro Gly Gly Pro Gln Trp Val Pro Gly Gly Ala Arg Gly Tyr
 85 90 95

<210> 235

<211> 404

<212> PRT

<213> Homo sapiens

<400> 235

Ile Gln Gln Trp Gly Asp Ser Val Leu Gly Arg Arg Cys Arg Asp Leu
 1 5 10 15

Leu Leu Gln Leu Tyr Leu Gln Arg Pro Glu Leu Arg Val Pro Val Pro
 20 25 30

Glu Val Leu Leu His Ser Glu Gly Ala Ala Ser Ser Ser Val Cys Lys
 35 40 45

Leu Asp Gly Leu Ile His Arg Phe Ile Thr Leu Leu Ala Asp Thr Ser
 50 55 60

Asp Ser Arg Ala Leu Glu Asn Arg Gly Ala Asp Ala Ser Met Ala Cys
 65 70 75 80

Arg Lys Leu Ala Val Ala His Pro Leu Leu Leu Leu Arg His Leu Pro
 85 90 95

Met Ile Ala Ala Leu Leu His Gly Arg Thr His Leu Asn Phe Gln Glu
 100 105 110

162

Phe Arg Gln Gln Asn His Leu Ser Cys Phe Leu His Val Leu Gly Leu
115 120 125

Leu Glu Leu Leu Gln Pro His Val Phe Arg Ser Glu His Gln Gly Ala
130 135 140

Leu Trp Asp Cys Leu Leu Ser Phe Ile Arg Leu Leu Leu Asn Tyr Arg
145 150 155 160

Lys Ser Ser Arg His Leu Ala Ala Phe Ile Asn Lys Phe Val Gln Phe
165 170 175

Ile His Lys Tyr Ile Thr Tyr Asn Ala Pro Ala Ala Ile Ser Phe Leu
180 185 190

Gln Lys His Ala Asp Pro Leu His Asp Leu Ser Phe Asp Asn Ser Asp
195 200 205

Leu Val Met Leu Lys Ser Leu Leu Ala Gly Leu Ser Leu Pro Ser Arg
210 215 220

Asp Asp Arg Thr Asp Arg Gly Leu Asp Glu Glu Gly Glu Glu Ser
225 230 235 240

Ser Ala Gly Ser Leu Pro Leu Val Ser Val Ser Leu Phe Thr Pro Leu
245 250 255

Thr Ala Ala Glu Met Ala Pro Tyr Met Lys Arg Leu Ser Arg Gly Gln
260 265 270

Thr Val Glu Asp Leu Leu Glu Val Leu Ser Asp Ile Asp Glu Met Ser
275 280 285

Arg Arg Arg Pro Glu Ile Leu Ser Phe Phe Ser Thr Asn Leu Gln Arg
290 295 300

Leu Met Ser Ser Ala Glu Glu Cys Cys Arg Asn Leu Ala Phe Ser Leu
305 310 315 320

Ala Leu Arg Ser Met Gln Asn Ser Pro Ser Ile Ala Ala Ala Phe Leu
325 330 335

Pro Thr Phe Met Tyr Cys Leu Gly Ser Gln Asp Phe Glu Val Val Gln
340 345 350

Thr Ala Leu Arg Asn Leu Pro Glu Tyr Ala Leu Leu Cys Gln Glu His
355 360 365

Ala Ala Val Leu Leu His Arg Ala Phe Leu Val Gly Met Tyr Gly Gln
370 375 380

Met Asp Pro Ser Ala Gln Ile Ser Glu Ala Leu Arg Ile Leu His Met
385 390 395 400

Glu Ala Val Met

163

<210> 236

<211> 361

<212> PRT

<213> Homo sapiens

<400> 236

Met Leu Leu Lys His Leu Gln Arg Met Val Ser Val Pro Gln Val Lys
1 5 10 15

Ala Ser Ala Leu Lys Val Val Thr Leu Thr Ala Asn Asp Lys Thr Ser
20 25 30

Val Ser Phe Ser Ser Leu Pro Gly Gln Gly Val Ile Tyr Asn Val Ile
35 40 45

Val Trp Asp Pro Phe Leu Asn Thr Ser Ala Ala Tyr Ile Pro Ala His
50 55 60

Thr Tyr Ala Cys Ser Phe Glu Ala Gly Glu Gly Ser Cys Ala Ser Leu
65 70 75 80

Gly Arg Val Ser Ser Lys Val Phe Phe Thr Leu Phe Ala Leu Leu Gly
85 90 95

Phe Phe Ile Cys Phe Phe Gly His Arg Phe Trp Lys Thr Glu Leu Phe
100 105 110

Phe Ile Gly Phe Ile Ile Met Gly Phe Phe Phe Tyr Ile Leu Ile Thr
115 120 125

Arg Leu Thr Pro Ile Lys Tyr Asp Val Asn Leu Ile Leu Thr Ala Val
130 135 140

Thr Gly Ser Val Gly Gly Met Phe Leu Val Ala Val Trp Trp Arg Phe
145 150 155 160

Gly Ile Leu Ser Ile Cys Met Leu Cys Val Gly Leu Val Leu Gly Phe
165 170 175

Leu Ile Ser Ser Val Thr Phe Phe Thr Pro Leu Gly Asn Leu Lys Ile
180 185 190

Phe His Asp Asp Gly Val Phe Trp Val Thr Phe Ser Cys Ile Ala Ile
195 200 205

Leu Ile Pro Val Val Phe Met Gly Cys Leu Arg Ile Leu Asn Ile Leu
210 215 220

Thr Cys Gly Val Ile Gly Ser Tyr Ser Val Val Leu Ala Ile Asp Ser
225 230 235 240

Tyr Trp Ser Thr Ser Leu Ser Tyr Ile Thr Leu Asn Val Leu Lys Arg
245 250 255

Ala Leu Asn Lys Asp Phe His Arg Ala Phe Thr Asn Val Pro Phe Gln
260 265 270

164

Thr Asn Asp Phe Ile Ile Leu Ala Val Trp Gly Met Leu Ala Val Ser
 275 280 285

Gly Ile Thr Leu Gln Ile Arg Arg Glu Arg Gly Arg Pro Phe Phe Pro
 290 295 300

Pro His Pro Tyr Lys Leu Trp Lys Gln Glu Arg Glu Arg Arg Val Thr
 305 310 315 320

Asn Ile Leu Asp Pro Ser Tyr His Ile Pro Pro Leu Arg Glu Arg Leu
 325 330 335

Tyr Gly Arg Leu Thr Gln Ile Lys Gly Leu Phe Gln Lys Glu Gln Pro
 340 345 350

Ala Gly Glu Arg Thr Pro Leu Leu Leu
 355 360

<210> 237

<211> 116

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (37)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (40)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 237

Trp Ala Arg Leu Arg Gly Pro Gly Ala His Ala Arg Thr Ser Pro Gln
 1 5 10 15

Pro Trp Arg Gly Pro Ser Pro Ala Gln Ala Ala Met Gly Phe Leu Gln
 20 25 30

Leu Leu Val Val Xaa Val Leu Xaa Ser Glu His Arg Val Ala Gly Ala
 35 40 45

Ala Glu Val Phe Gly Asn Ser Ser Glu Gly Leu Ile Glu Phe Ser Val
 50 55 60

Gly Lys Phe Arg Tyr Phe Glu Leu Asn Arg Pro Phe Pro Glu Glu Ala
 65 70 75 80

Ile Leu His Asp Ile Ser Ser Asn Val Thr Phe Leu Ile Phe Gln Ile
 85 90 95

His Ser Gln Tyr Gln Asn Thr Thr Val Ser Phe Ser Pro Arg Arg Arg
 100 105 110

Ser Pro Thr Met
 115

165

<210> 238
 <211> 166
 <212> PRT
 <213> Homo sapiens

<400> 238
 Pro Arg Val Arg Pro Ala Ser Pro Pro Val Arg Ser Pro Ala Arg Trp
 1 5 10 15
 Gly Ser Met Ala Gly Ser Pro Leu Leu Trp Gly Pro Arg Ala Gly Gly
 20 25 30
 Val Gly Leu Leu Val Leu Leu Leu Leu Gly Leu Phe Arg Pro Pro Pro
 35 40 45
 Ala Leu Cys Ala Arg Pro Val Lys Glu Pro Arg Gly Leu Ser Ala Ala
 50 55 60
 Ser Pro Pro Leu Ala Arg Leu Ala Leu Leu Ala Ala Ser Gly Gly Gln
 65 70 75 80
 Cys Pro Glu Val Arg Arg Arg Gly Arg Cys Arg Pro Gly Ala Gly Ala
 85 90 95
 Gly Ala Ser Ala Gly Ala Glu Arg Gln Glu Arg Ala Arg Ala Glu Ala
 100 105 110
 Gln Arg Leu Arg Ile Ser Arg Arg Ala Ser Trp Arg Ser Cys Cys Ala
 115 120 125
 Ser Gly Ala Pro Pro Ala Thr Leu Ile Arg Leu Trp Ala Trp Thr Thr
 130 135 140
 Thr Pro Thr Arg Leu Gln Arg Ser Ser Leu Ala Leu Cys Ser Ala Pro
 145 150 155 160
 Ala Leu Thr Leu Pro Pro
 165

<210> 239
 <211> 414
 <212> PRT
 <213> Homo sapiens

<400> 239
 Pro Arg Val Arg Leu Ala Thr Pro Asn Ile Trp Asp Leu Ser Met Leu
 1 5 10 15
 Phe Ala Phe Ile Ser Leu Leu Val Met Leu Pro Thr Trp Trp Ile Val
 20 25 30
 Ser Ser Trp Leu Val Trp Gly Val Ile Leu Phe Val Tyr Leu Val Ile
 35 40 45
 Arg Ala Leu Arg Leu Trp Arg Thr Ala Lys Leu Gln Val Thr Leu Lys

166

50	55	60
Lys Tyr Ser Val His Leu Glu Asp Met Ala Thr Asn Ser Arg Ala Phe		
65	70	75
Thr Asn Leu Val Arg Lys Ala Leu Arg Leu Ile Gln Glu Thr Glu Val		
	85	90
Ile Ser Arg Gly Phe Thr Leu Val Ser Ala Ala Cys Pro Phe Asn Lys		
	100	105
Ala Gly Gln His Pro Ser Gln His Leu Ile Gly Leu Arg Lys Ala Val		
	115	120
Tyr Arg Thr Leu Arg Ala Asn Phe Gln Ala Ala Arg Leu Ala Thr Leu		
	130	135
Tyr Met Leu Lys Asn Tyr Pro Leu Asn Ser Glu Ser Asp Asn Val Thr		
	145	150
Asn Tyr Ile Cys Val Val Pro Phe Lys Glu Leu Gly Leu Gly Leu Ser		
	165	170
Glu Glu Gln Ile Ser Glu Glu Glu Ala His Asn Phe Thr Asp Gly Phe		
	180	185
Ser Leu Pro Ala Leu Lys Val Leu Phe Gln Leu Trp Val Ala Gln Ser		
	195	200
Ser Glu Phe Phe Arg Arg Leu Ala Leu Leu Leu Ser Thr Ala Asn Ser		
	210	215
Pro Pro Gly Pro Leu Leu Thr Pro Ala Leu Leu Pro His Arg Ile Leu		
	225	230
Ser Asp Val Thr Gln Gly Leu Pro His Ala His Ser Ala Cys Leu Glu		
	245	250
Glu Leu Lys Arg Ser Tyr Glu Phe Tyr Arg Tyr Phe Glu Thr Gln His		
	260	265
Gln Ser Val Pro Gln Cys Leu Ser Lys Thr Gln Gln Lys Ser Arg Glu		
	275	280
Leu Asn Asn Val His Thr Ala Val Arg Ser Leu Gln Leu His Leu Lys		
	290	295
Ala Leu Leu Asn Glu Val Ile Ile Leu Glu Asp Glu Leu Glu Lys Leu		
	305	310
Val Cys Thr Lys Glu Thr Gln Glu Leu Val Ser Glu Ala Tyr Pro Ile		
	325	330
Leu Glu Gln Lys Leu Lys Leu Ile Gln Pro His Val Gln Ala Ser Asn		
	340	345
Asn Cys Trp Glu Glu Ala Ile Ser Gln Val Asp Lys Leu Leu Arg Arg		
	355	360

167

Asn Thr Asp Lys Lys Gly Lys Pro Glu Ile Ala Cys Glu Asn Pro His
370 375 380

Cys Thr Val Ser Thr Phe Glu Ala Ala Tyr Ser Thr His Cys Arg Gln
385 390 395 400

Arg Ser Asn Pro Arg Gly Ala Gly Ile Arg Ser Leu Cys Arg
405 410

<210> 240

<211> 145

<212> PRT

<213> Homo sapiens

<400> 240

Ala Ala Pro His Pro Pro Leu Leu Arg Pro Leu Cys Leu Trp Cys Pro
1 5 10 15

Leu Trp Pro Ala Trp Pro Leu Arg Gly Arg Pro Arg Ser Ala Trp Lys
20 25 30

Arg Trp Pro Pro Leu Pro Val Gly Pro Ala Lys Leu Gly Cys Ser Met
35 40 45

Thr Thr Arg Gln Pro Thr Ala Val Ser Trp Pro Cys Trp Leu Met Ser
50 55 60

Ser Ser Leu Ser Thr Ala Cys Leu Ala Trp Thr Leu Thr Gly Ser Leu
65 70 75 80

Ala Arg Glu Ala Thr Arg Arg Ala Arg Ser Leu Ser Pro Thr Trp Asn
85 90 95

Cys Ser Ala Arg Gln Val Pro Pro Ser Pro Pro His Ser Gly Leu Gly
100 105 110

Arg Arg Gly Trp Ala His Cys His Leu Thr Cys Leu Leu Val Thr Gln
115 120 125

Leu Phe Arg Val Gly Arg Ile His Pro Ile Leu Ser Leu Pro Leu Val
130 135 140

Thr

145

<210> 241

<211> 72

<212> PRT

<213> Homo sapiens

<400> 241

Leu Gln Leu Ala Ser Gln Ser Ala Gly Ile Lys Gly Met Ser His Cys
1 5 10 15

Ala Arg Pro Thr Phe Leu Thr Leu Leu Leu Ala Ser Cys Phe Trp Ala

168

20 25 30

Ala Ala Ile Pro Asn Arg Asn Val Ile Leu Ser Val Ser Phe Arg Pro
35 40 45

Leu His Met Gln Phe Thr Leu Ser Ile Leu Val Phe Ile Leu Arg Ile
50 55 60

Leu Ile Leu Leu Arg Ser Phe Leu
65 70

<210> 242
<211> 140
<212> PRT
<213> Homo sapiens

<400> 242
Met Val Leu Val Leu Arg His Pro Leu Cys Ala Arg Glu Arg Ala Phe
1 5 10 15

Arg Glu Pro Gly Arg Gly Leu Leu Thr Arg Thr Gly Gln His Asp Gly
20 25 30

Ala Pro Ala Val Thr Ala Val Pro Gly Pro Leu Gly Ala Val Ala Ala
35 40 45

Ala Glu Gly Arg Arg Ser Ala Trp Gly Ala Gly Gly Ser Ser Pro Pro
50 55 60

Arg Lys Val Leu Trp Gly Asp Met Arg Gly Arg Arg Ala Gly Val Asp
65 70 75 80

Val Leu Gly Pro Ala Leu Ser Ser Glu Ala Ala Gly Ala Glu Ala Arg
85 90 95

Gly Trp Gly Met Pro Gly Met Gly Val Gly Val Gly Ala Ser Glu Thr
100 105 110

Arg Gly Ala Leu Phe Leu Gly Arg Glu Gly Val His Gly Pro Cys Pro
115 120 125

Met Asp Gly Leu Gly Pro Trp Pro Trp Gly Pro Trp
130 135 140

<210> 243
<211> 353
<212> PRT
<213> Homo sapiens

<400> 243
Met Gly Pro Ala Val Lys Met Trp Thr Asn Ala Trp Lys Gly Leu Asp
1 5 10 15

Asp Cys His Tyr Asn Gln Leu Cys Glu Asn Thr Pro Gly Gly His Arg
20 25 30

169

Cys Ser Cys Pro Arg Gly Tyr Arg Met Gln Gly Pro Ser Leu Pro Cys
 35 40 45
 Leu Asp Val Asn Glu Cys Leu Gln Leu Pro Lys Ala Cys Ala Tyr Gln
 50 55 60
 Cys His Asn Leu Gln Gly Ser Tyr Arg Cys Leu Cys Pro Pro Gly Gln
 65 70 75 80
 Thr Leu Leu Arg Asp Gly Lys Ala Cys Thr Ser Leu Glu Arg Asn Gly
 85 90 95
 Gln Asn Val Thr Thr Val Ser His Arg Gly Pro Leu Leu Pro Trp Leu
 100 105 110
 Arg Pro Trp Ala Ser Ile Pro Gly Thr Ser Tyr His Ala Trp Val Ser
 115 120 125
 Leu Arg Pro Gly Pro Met Ala Leu Ser Ser Val Gly Arg Ala Trp Cys
 130 135 140
 Pro Pro Gly Phe Ile Arg Gln Asn Gly Val Cys Thr Asp Leu Asp Glu
 145 150 155 160
 Cys Arg Val Arg Asn Leu Cys Gln His Ala Cys Arg Asn Thr Glu Gly
 165 170 175
 Ser Tyr Gln Cys Leu Cys Pro Ala Gly Tyr Arg Leu Leu Pro Ser Gly
 180 185 190
 Lys Asn Cys Gln Asp Ile Asn Glu Cys Glu Glu Glu Ser Ile Glu Cys
 195 200 205
 Gly Pro Gly Gln Met Cys Phe Asn Thr Arg Gly Ser Tyr Gln Cys Val
 210 215 220
 Asp Thr Pro Cys Pro Ala Thr Tyr Arg Gln Gly Pro Ser Pro Gly Thr
 225 230 235 240
 Cys Phe Arg Arg Cys Ser Gln Asp Cys Gly Thr Gly Gly Pro Ser Thr
 245 250 255
 Leu Gln Tyr Arg Leu Leu Pro Leu Pro Leu Gly Val Arg Ala His His
 260 265 270
 Asp Val Ala Arg Leu Thr Ala Phe Ser Glu Val Gly Val Pro Ala Asn
 275 280 285
 Arg Thr Glu Leu Ser Met Leu Glu Pro Asp Pro Arg Ser Pro Phe Ala
 290 295 300
 Leu Arg Pro Leu Arg Ala Gly Leu Gly Ala Val Tyr Thr Arg Arg Ala
 305 310 315 320
 Leu Thr Arg Ala Gly Leu Tyr Arg Leu Thr Val Arg Ala Ala Ala Pro
 325 330 335
 Arg His Gln Ser Val Phe Val Leu Leu Ile Ala Val Ser Pro Tyr Pro

170

340

345

350

Tyr

<210> 244

<211> 146

<212> PRT

<213> Homo sapiens

<400> 244

Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala
 1 5 10 15

Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys
 20 25 30

Phe Arg Gln Ser His Thr Trp Ser Leu Leu Gln Gly Ala Ala His Asp
 35 40 45

Asp Pro Val Ala Arg Gly Leu Asp Pro Asp Gly Leu Leu Leu Asp
 50 55 60

Val Val Val Asn Gly Val Val Pro Gly Arg Ala Trp Leu Thr Gln Ile
 65 70 75 80

Phe Lys Cys Arg Thr Leu Lys Lys His Tyr Val Gln Thr Arg Ala Trp
 85 90 95

Pro Ala Val Arg Gly Leu His Thr Ala Leu Leu Pro Gly Arg Pro Pro
 100 105 110

Leu Val Pro Thr Leu Gln Pro Gln His Pro Val Gln Arg Gly Pro Gly
 115 120 125

Pro Pro Ala Pro Ala Gly Ala Ala Pro Ala Gly Leu Ser Tyr Gln Leu
 130 135 140

Gly Leu
 145

<210> 245

<211> 638

<212> PRT

<213> Homo sapiens

<400> 245

His Ala Ser Gly Ala Phe Leu Val Val Arg Gly Glu Pro Gln Gly Ser
 1 5 10 15

Trp Gly Ser Met Thr Gly Val Ile Asn Gly Arg Lys Phe Gly Val Ala
 20 25 30

Thr Leu Asn Thr Ser Val Met Gln Glu Ala His Ser Gly Val Ser Ser
 35 40 45

171

Ile His Ser Ser Ile Arg His Val Pro Ala Asn Val Gly Pro Leu Met
 50 55 60
 Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala Arg
 65 70 75 80
 Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys Phe
 85 90 95
 Arg Gln Glu Ser His Val Glu Phe Ala Thr Gly Glu Leu Leu Thr Met
 100 105 110
 Thr Gln Trp Pro Gly Val Trp Ile Pro Met Ala Ser Cys Ser Ser Thr
 115 120 125
 Trp Trp Ser Met Ala Leu Ser Pro Asp Ser Leu Ala Asp Ala Asp Leu
 130 135 140
 Gln Val Gln Asp Phe Glu Glu His Tyr Val Gln Thr Gly Pro Gly Gln
 145 150 155 160
 Leu Phe Val Gly Ser Thr Gln Arg Phe Phe Gln Gly Gly Leu Pro Ser
 165 170 175
 Phe Leu Arg Cys Asn His Ser Ile Gln Tyr Asn Ala Ala Arg Gly Pro
 180 185 190
 Gln Pro Gln Leu Val Gln His Leu Arg Ala Ser Ala Ile Ser Ser Ala
 195 200 205
 Phe Asp Pro Glu Ala Glu Ala Leu Arg Phe Gln Leu Ala Thr Ala Leu
 210 215 220
 Gln Ala Glu Glu Asn Glu Val Gly Cys Pro Glu Gly Phe Glu Leu Asp
 225 230 235 240
 Ser Gln Gly Ala Phe Cys Val Asp Val Asp Glu Cys Ala Trp Asp Ala
 245 250 255
 His Leu Cys Arg Glu Gly Gln Arg Cys Val Asn Leu Leu Gly Ser Tyr
 260 265 270
 Arg Cys Leu Pro Asp Cys Gly Pro Gly Phe Arg Val Ala Asp Gly Ala
 275 280 285
 Gly Cys Glu Asp Val Asp Glu Cys Leu Glu Gly Leu Asp Asp Cys His
 290 295 300
 Tyr Asn Gln Leu Cys Glu Asn Thr Pro Gly Gly His Arg Cys Ser Cys
 305 310 315 320
 Pro Arg Gly Tyr Arg Met Gln Gly Pro Ser Leu Pro Cys Leu Asp Val
 325 330 335
 Asn Glu Cys Leu Gln Leu Pro Lys Ala Cys Ala Tyr Gln Cys His Asn
 340 345 350
 Leu Gln Gly Ser Tyr Arg Cys Leu Cys Pro Pro Gly Gln Thr Leu Leu

172

355 360 365
 Arg Asp Gly Lys Ala Cys Thr Ser Leu Glu Arg Asn Gly Gln Asn Val
 370 375 380
 Thr Thr Val Ser His Arg Gly Pro Leu Leu Pro Trp Leu Arg Pro Trp
 385 390 395 400
 Ala Ser Ile Pro Gly Thr Ser Tyr His Ala Trp Val Ser Leu Arg Pro
 405 410 415
 Gly Pro Met Ala Leu Ser Ser Val Gly Arg Ala Trp Cys Pro Pro Gly
 420 425 430
 Phe Ile Arg Gln Asn Gly Val Cys Thr Asp Leu Asp Glu Cys Arg Val
 435 440 445
 Arg Asn Leu Cys Gln His Ala Cys Arg Asn Thr Glu Gly Ser Tyr Gln
 450 455 460
 Cys Leu Cys Pro Ala Gly Tyr Arg Leu Leu Pro Ser Gly Lys Asn Cys
 465 470 475 480
 Gln Asp Ile Asn Glu Cys Glu Glu Glu Ser Ile Glu Cys Gly Pro Gly
 485 490 495
 Gln Met Cys Phe Asn Thr Arg Gly Ser Tyr Gln Cys Val Asp Thr Pro
 500 505 510
 Cys Pro Ala Thr Tyr Arg Gln Gly Pro Ser Pro Gly Thr Cys Phe Arg
 515 520 525
 Arg Cys Ser Gln Asp Cys Gly Thr Gly Gly Pro Ser Thr Leu Gln Tyr
 530 535 540
 Arg Leu Leu Pro Leu Pro Leu Gly Val Arg Ala His His Asp Val Ala
 545 550 555 560
 Arg Leu Thr Ala Phe Ser Glu Val Gly Val Pro Ala Asn Arg Thr Glu
 565 570 575
 Leu Ser Met Leu Glu Pro Asp Pro Arg Ser Pro Phe Ala Leu Arg Pro
 580 585 590
 Leu Arg Ala Gly Leu Gly Ala Val Tyr Thr Arg Arg Ala Leu Thr Arg
 595 600 605
 Ala Gly Leu Tyr Arg Leu Thr Val Arg Ala Ala Ala Pro Arg His Gln
 610 615 620
 Ser Val Phe Val Leu Leu Ile Ala Val Ser Pro Tyr Pro Tyr
 625 630 635

<210> 246

<211> 367

<212> PRT

<213> Homo sapiens

173

<400> 246

```

Met Gly Glu Lys Phe Leu Leu Leu Ala Met Lys Glu Asn His Pro Glu
 1              5              10              15

Cys Phe Cys Lys Ile Leu Lys Ile Leu His Cys Met Asp Pro Gly Glu
          20              25              30

Trp Leu Pro Gln Thr Glu His Cys Val His Leu Thr Pro Lys Glu Phe
          35              40              45

Leu Ile Trp Thr Met Asp Ile Ala Ser Asn Glu Arg Ser Glu Ile Gln
 50              55              60

Ser Val Ala Leu Arg Leu Ala Ser Lys Val Ile Ser His His Met Gln
 65              70              75              80

Thr Cys Val Glu Asn Arg Glu Leu Ile Ala Ala Glu Leu Lys Gln Trp
          85              90              95

Val Gln Leu Val Ile Leu Ser Cys Glu Asp His Leu Pro Thr Glu Ser
          100              105              110

Arg Leu Ala Val Val Glu Val Leu Thr Ser Thr Thr Pro Leu Phe Leu
          115              120              125

Thr Asn Pro His Pro Ile Leu Glu Leu Gln Asp Thr Leu Ala Leu Trp
          130              135              140

Lys Cys Val Leu Thr Leu Leu Gln Ser Glu Glu Gln Ala Val Arg Asp
          145              150              155              160

Ala Ala Thr Glu Thr Val Thr Thr Ala Met Ser Gln Glu Asn Thr Cys
          165              170              175

Gln Ser Thr Glu Phe Ala Phe Cys Gln Val Asp Ala Ser Ile Ala Leu
          180              185              190

Ala Leu Ala Leu Ala Val Leu Cys Asp Leu Leu Gln Gln Trp Asp Gln
          195              200              205

Leu Ala Pro Gly Leu Pro Ile Leu Leu Gly Trp Leu Leu Gly Glu Ser
          210              215              220

Asp Asp Leu Val Ala Cys Val Glu Ser Met His Gln Val Glu Glu Asp
          225              230              235              240

Tyr Leu Phe Glu Lys Ala Glu Val Asn Phe Trp Ala Glu Thr Leu Ile
          245              250              255

Phe Val Lys Tyr Leu Cys Lys His Leu Phe Cys Leu Leu Ser Lys Ser
          260              265              270

Gly Trp Arg Pro Pro Ser Pro Glu Met Leu Cys His Leu Gln Arg Met
          275              280              285

Val Ser Glu Gln Cys His Leu Leu Ser Gln Phe Phe Arg Glu Leu Pro
          290              295              300

```

174

Pro Ala Ala Glu Phe Val Lys Thr Val Glu Phe Thr Arg Leu Arg Ile
 305 310 315 320

Gln Glu Glu Arg Thr Leu Ala Cys Leu Arg Leu Leu Ala Phe Leu Glu
 325 330 335

Gly Lys Glu Gly Glu Asp Thr Leu Val Leu Ser Val Trp Asp Ser Tyr
 340 345 350

Ala Glu Ser Arg Gln Leu Thr Leu Pro Arg Thr Glu Ala Ala Cys
 355 360 365

<210> 247
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 247
 Met Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Leu Val
 1 5 10 15

Leu Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu
 20 25 30

Ser Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro
 35 40 45

Val Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val
 50 55 60

Met Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu
 65 70 75 80

Pro Ser Cys Thr Asp Gln Cys Phe Arg Ala Lys Pro His Ser Trp Gly
 85 90 95

His Phe Ser Arg Phe Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys
 100 105 110

Thr Arg Asn Glu Phe Arg Leu Pro Ala Arg Ala Asp
 115 120

<210> 248
 <211> 674
 <212> PRT
 <213> Homo sapiens

<400> 248
 Met Thr Gly Arg Glu Phe Phe Ser Arg Phe Pro Glu Leu Tyr Pro Phe
 1 5 10 15

Leu Leu Lys Gln Leu Glu Thr Val Ala Asn Thr Val Asp Ser Asp Met
 20 25 30

Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Val Leu

175

35	40	45
Glu Arg Leu Tyr Ala Ser	Pro Met Asp Gly Thr Ser	Ser Ala Leu Ser
50	55	60
Met Gly Pro Phe Val	Pro Phe Ile Met Arg Cys Gly His Ser	Pro Val
65	70	75 80
Tyr His Ser Arg Glu Met	Ala Ala Arg Ala Leu Val	Pro Phe Val Met
85	90	95
Ile Asp His Ile Pro Asn Thr	Ile Arg Thr Leu Leu Ser	Thr Leu Pro
100	105	110
Ser Cys Thr Asp Gln Cys Phe	Arg Gln Asn His Ile His Gly Thr Leu	
115	120	125
Leu Gln Val Phe His Leu Leu	Gln Ala Tyr Ser Asp Ser Lys His Gly	
130	135	140
Thr Asn Ser Asp Phe Gln His Glu Leu Thr	Asp Ile Thr Val Cys Thr	
145	150	155 160
Lys Ala Lys Leu Trp Leu Ala Lys Arg	Gln Asn Pro Cys Leu Val Thr	
165	170	175
Arg Ala Val Tyr Ile Asp Ile Leu Phe	Leu Leu Thr Cys Cys Leu Asn	
180	185	190
Arg Ser Ala Lys Asp Asn Gln Pro Val	Leu Glu Ser Leu Gly Phe Trp	
195	200	205
Glu Glu Val Arg Gly Ile Ile Ser Gly	Ser Glu Leu Ile Thr Gly Phe	
210	215	220
Pro Trp Ala Phe Lys Val Pro Gly Leu Pro	Gln Tyr Leu Gln Ser Leu	
225	230	235 240
Thr Arg Leu Ala Ile Ala Ala Val Trp	Ala Ala Ala Ala Lys Ser Gly	
245	250	255
Glu Arg Glu Thr Asn Val Pro Ile Ser Phe	Ser Gln Leu Leu Glu Ser	
260	265	270
Ala Phe Pro Glu Val Arg Ser Leu Thr	Leu Glu Ala Leu Leu Glu Lys	
275	280	285
Phe Leu Ala Ala Ala Ser Gly Leu Gly	Glu Lys Gly Val Pro Pro Leu	
290	295	300
Leu Cys Asn Met Gly Glu Lys Phe Leu Leu	Leu Ala Met Lys Glu Asn	
305	310	315 320
His Pro Glu Cys Phe Cys Lys Ile Leu Lys	Ile Leu His Cys Met Asp	
325	330	335
Pro Gly Glu Trp Leu Pro Gln Thr Glu	His Cys Val His Leu Thr Pro	
340	345	350

176

Lys Glu Phe Leu Ile Trp Thr Met Asp Ile Ala Ser Asn Glu Arg Ser
 355 360 365
 Glu Ile Gln Ser Val Ala Leu Arg Leu Ala Ser Lys Val Ile Ser His
 370 375 380
 His Met Gln Thr Cys Val Glu Asn Arg Glu Leu Ile Ala Ala Glu Leu
 385 390 395 400
 Lys Gln Trp Val Gln Leu Val Ile Leu Ser Cys Glu Asp His Leu Pro
 405 410 415
 Thr Glu Ser Arg Leu Ala Val Val Glu Val Leu Thr Ser Thr Thr Pro
 420 425 430
 Leu Phe Leu Thr Asn Pro His Pro Ile Leu Glu Leu Gln Asp Thr Leu
 435 440 445
 Ala Leu Trp Lys Cys Val Leu Thr Leu Leu Gln Ser Glu Glu Gln Ala
 450 455 460
 Val Arg Asp Ala Ala Thr Glu Thr Val Thr Thr Ala Met Ser Gln Glu
 465 470 475 480
 Asn Thr Cys Gln Ser Thr Glu Phe Ala Phe Cys Gln Val Asp Ala Ser
 485 490 495
 Ile Ala Leu Ala Leu Ala Leu Ala Val Leu Cys Asp Leu Leu Gln Gln
 500 505 510
 Trp Asp Gln Leu Ala Pro Gly Leu Pro Ile Leu Leu Gly Trp Leu Leu
 515 520 525
 Gly Glu Ser Asp Asp Leu Val Ala Cys Val Glu Ser Met His Gln Val
 530 535 540
 Glu Glu Asp Tyr Leu Phe Glu Lys Ala Glu Val Asn Phe Trp Ala Glu
 545 550 555 560
 Thr Leu Ile Phe Val Lys Tyr Leu Cys Lys His Leu Phe Cys Leu Leu
 565 570 575
 Ser Lys Ser Gly Trp Arg Pro Pro Ser Pro Glu Met Leu Cys His Leu
 580 585 590
 Gln Arg Met Val Ser Glu Gln Cys His Leu Leu Ser Gln Phe Phe Arg
 595 600 605
 Glu Leu Pro Pro Ala Ala Glu Phe Val Lys Thr Val Glu Phe Thr Arg
 610 615 620
 Leu Arg Ile Gln Glu Glu Arg Thr Leu Ala Cys Leu Arg Leu Leu Ala
 625 630 635 640
 Phe Leu Glu Gly Lys Glu Gly Glu Asp Thr Leu Val Leu Ser Val Trp
 645 650 655

177

Asp Ser Tyr Ala Glu Ser Arg Gln Leu Thr Leu Pro Arg Thr Glu Ala
 660 665 670

Ala Cys

<210> 249
 <211> 10
 <212> PRT
 <213> Homo sapiens

<400> 249
 Ile Ile Ser Gly Ser Glu Leu Ile Thr Gly
 1 5 10

<210> 250
 <211> 230
 <212> PRT
 <213> Homo sapiens

<400> 250
 Val Asp Gly Ile Asp Lys Leu Asp Ile Glu Phe Leu Gln Gln Phe Leu
 1 5 10 15

Glu Thr His Ser Arg Gly Pro Arg Leu His Ser Pro Gly His Ala Ser
 20 25 30

Gln Glu Ala Thr Pro Gly Ala Asn Met Ser Ser Gly Thr Glu Leu Leu
 35 40 45

Trp Pro Gly Ala Ala Leu Leu Val Leu Leu Gly Val Ala Ala Ser Leu
 50 55 60

Cys Val Arg Cys Ser Arg Pro Gly Ala Lys Arg Ser Glu Lys Ile Tyr
 65 70 75 80

Gln Gln Arg Ser Leu Arg Glu Asp Gln Gln Ser Phe Thr Gly Ser Arg
 85 90 95

Thr Tyr Ser Leu Val Gly Gln Ala Trp Pro Gly Pro Leu Ala Asp Met
 100 105 110

Ala Pro Thr Arg Lys Asp Lys Leu Leu Gln Phe Tyr Pro Ser Leu Glu
 115 120 125

Asp Pro Ala Ser Ser Arg Tyr Gln Asn Phe Ser Lys Gly Ser Arg His
 130 135 140

Gly Ser Glu Glu Ala Tyr Ile Asp Pro Ile Ala Met Glu Tyr Tyr Asn
 145 150 155 160

Trp Gly Arg Phe Ser Lys Pro Pro Glu Asp Asp Asp Ala Asn Ser Tyr
 165 170 175

Glu Asn Val Leu Ile Cys Lys Gln Lys Thr Thr Glu Thr Gly Ala Gln
 180 185 190

178

Gln Glu Gly Ile Gly Gly Leu Cys Arg Gly Asp Leu Ser Leu Ser Leu
195 200 205

Ala Leu Lys Thr Gly Pro Thr Ser Gly Leu Cys Pro Ser Ala Ser Pro
210 215 220

Glu Glu Asp Glu Gly Ile
225 230

<210> 251
<211> 122
<212> PRT
<213> Homo sapiens

<400> 251
Val Leu Trp Arg Glu Ala Ser Ala Leu Val Leu Ser Asn Arg Leu Ser
1 5 10 15

Ser Gly Leu Leu His Asp Leu Leu Leu Gln Pro Ala Ile His Ser Arg
20 25 30

Leu Phe Pro Arg Arg Ser Arg Gly Leu Ser Glu Gly Glu Gly Ser Ser
35 40 45

Val Ser Leu Gln Arg Ser Arg Val Leu Ser Ala Met Lys His Val Leu
50 55 60

Asn Leu Tyr Leu Leu Gly Val Val Leu Thr Leu Leu Ser Ile Phe Val
65 70 75 80

Arg Val Met Glu Ser Leu Glu Gly Leu Leu Glu Ser Pro Ser Pro Gly
85 90 95

Thr Ser Trp Thr Thr Arg Ser Gln Leu Ala Asn Thr Glu Pro Thr Lys
100 105 110

Gly Leu Pro Asp His Pro Ser Arg Ser Met
115 120

<210> 252
<211> 129
<212> PRT
<213> Homo sapiens

<400> 252
Tyr Thr Phe His Thr Gln Ile Phe Leu Asp Phe Pro Met Ile Phe Leu
1 5 10 15

Thr Val Leu Pro Leu Ala Phe Leu Phe Leu His Ser Gly Phe Tyr His
20 25 30

Tyr Ile Ser Phe Ser Cys Leu Phe Ser Leu Ser Leu Ala Leu Phe Phe
35 40 45

Phe Leu Asp Val Ala Thr Phe Arg Arg Pro Gly Gln Leu Phe Cys Glu

179

50 55 60
 Arg Ser Val Leu Phe Asp Met Phe His Phe Gly Phe Val Ser Leu Phe
 65 70 75 80
 Leu His Glu Trp Ile Gln Ala Lys His Phe Trp Ala Gly Leu Phe Ile
 85 90 95
 Val Leu Pro Ser Asp Val Phe Phe Ser Val His His Leu Glu Ala Pro
 100 105 110
 Asp Gly Ser Phe Pro Asn Ile Ala Lys Leu Ser Leu Ile Ile Leu Leu
 115 120 125

Arg

<210> 253
 <211> 99
 <212> PRT
 <213> Homo sapiens

<400> 253
 Gly Thr Arg Phe Pro Thr Gly Glu Thr Pro Ser Leu Gly Phe Thr Val
 1 5 10 15
 Thr Leu Val Leu Leu Asn Ser Leu Ala Phe Leu Leu Met Ala Val Ile
 20 25 30
 Tyr Thr Lys Leu Tyr Cys Asn Leu Glu Lys Glu Asp Leu Ser Glu Asn
 35 40 45
 Ser Gln Ser Ser Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn
 50 55 60
 Cys Ile Phe Phe Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile
 65 70 75 80
 Thr Ala Ile Ser Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile
 85 90 95
 Phe Phe Pro

<210> 254
 <211> 51
 <212> PRT
 <213> Homo sapiens

<400> 254
 Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn Cys Ile Phe Phe
 1 5 10 15
 Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile Thr Ala Ile Ser
 20 25 30

180

Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile Phe Phe Pro Cys
 35 40 45

Leu Leu Ala
 50

<210> 255

<211> 259

<212> PRT

<213> Homo sapiens

<400> 255

Gly Thr Arg Phe Pro Thr Gly Glu Thr Pro Ser Leu Gly Phe Thr Val
 1 5 10 15

Thr Leu Val Leu Leu Asn Ser Leu Ala Phe Leu Leu Met Ala Val Ile
 20 25 30

Tyr Thr Lys Leu Tyr Cys Asn Leu Glu Lys Glu Asp Leu Ser Glu Asn
 35 40 45

Ser Gln Ser Ser Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn
 50 55 60

Cys Ile Phe Phe Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile
 65 70 75 80

Thr Ala Ile Ser Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile
 85 90 95

Phe Phe Pro Leu Pro Ala Cys Leu Asn Pro Val Leu Tyr Val Phe Phe
 100 105 110

Asn Pro Lys Phe Lys Glu Asp Trp Lys Leu Leu Lys Arg Arg Val Thr
 115 120 125

Lys Lys Ser Gly Ser Val Ser Val Ser Ile Ser Ser Gln Gly Gly Cys
 130 135 140

Leu Glu Gln Asp Phe Tyr Tyr Asp Cys Gly Met Tyr Ser His Leu Gln
 145 150 155 160

Gly Asn Leu Thr Val Cys Asp Cys Cys Glu Ser Phe Leu Leu Thr Lys
 165 170 175

Pro Val Ser Cys Lys His Leu Ile Lys Ser His Ser Cys Pro Ala Leu
 180 185 190

Ala Val Ala Ser Cys Gln Arg Pro Glu Gly Tyr Trp Ser Asp Cys Gly
 195 200 205

Thr Gln Ser Ala His Ser Asp Tyr Ala Asp Glu Glu Asp Ser Phe Val
 210 215 220

Ser Asp Ser Ser Asp Gln Val Gln Ala Cys Gly Arg Ala Cys Phe Tyr
 225 230 235 240

<400> 258
Gly His Glu Ser Ile Cys Gly Ser Cys Arg Ser Trp Ile Tyr Phe Ser
1 5 10 15

182

Ile Arg Cys Arg Arg Arg Met Arg Pro Trp Trp Ser Leu Leu Leu Glu
 20 25 30
 Ala Cys Ala Thr Cys Ala Gln Thr Gly Pro Thr Arg Ser Thr Ser Cys
 35 40 45
 Thr Gln Glu Val Ser His Ser Ser Ser Thr Ala Tyr Pro Ala Pro Met
 50 55 60
 Arg Arg Arg Cys Cys Leu Pro Ser Pro Arg Ser Cys Thr
 65 70 75

<210> 259
 <211> 119
 <212> PRT
 <213> Homo sapiens

<400> 259
 Lys Arg Ala Gly Val Glu Val Gly Gly Leu Val Met Ala Leu Ala Gly
 1 5 10 15
 Ser Val Phe Val Leu Gly Gly Val Leu Val Leu Cys Val Glu Arg Asn
 20 25 30
 Gly Glu Gly Glu Met Gly Trp Pro Gln His Leu Pro Lys Ser Gln Pro
 35 40 45
 Leu Ser Pro Pro Val Ala Val Arg Arg Cys Ser Phe Glu Arg Ser Trp
 50 55 60
 Ile Asp Leu Leu Val Glu Thr Ser Ser Ser Met Val Thr Cys Arg Gln
 65 70 75 80
 Gln Val Gly Thr Pro Asn Gly Met Glu Gly Arg Gly Gly Gly Pro Lys
 85 90 95
 Thr Thr Phe Pro Ile Arg Leu Gln Leu Ser Gly Ala Cys Ala Val Arg
 100 105 110
 Pro Glu Ile Gln Trp Glu Val
 115

<210> 260
 <211> 275
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (47)
 <223> Xaa equals any of the naturally occurring L-amino acids
 <220>
 <221> SITE
 <222> (94)

183

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (192)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 260

Gln	Asp	Trp	Lys	Ala	Glu	Arg	Ser	Gln	Asp	Pro	Phe	Glu	Lys	Cys	Met
1				5					10					15	

Gln	Asp	Pro	Asp	Tyr	Glu	Gln	Leu	Leu	Lys	Val	Thr	Ile	Leu	Glu	Ala
			20					25					30		

Asp	Asn	Arg	Ile	Gly	Gly	Arg	Ile	Phe	Thr	Tyr	Arg	Asp	Gln	Xaa	Thr
	35					40						45			

Gly	Trp	Ile	Gly	Glu	Leu	Gly	Ala	Met	Arg	Met	Pro	Ser	Ser	His	Arg
	50					55					60				

Ile	Leu	His	Lys	Leu	Cys	Gln	Gly	Leu	Gly	Leu	Asn	Leu	Thr	Lys	Phe
	65				70					75					80

Thr	Gln	Tyr	Asp	Lys	Asn	Thr	Trp	Thr	Glu	Val	His	Glu	Xaa	Lys	Leu
				85					90					95	

Arg	Asn	Tyr	Val	Val	Glu	Lys	Val	Pro	Glu	Lys	Leu	Gly	Tyr	Ala	Leu
			100						105				110		

Arg	Pro	Gln	Glu	Lys	Gly	His	Ser	Pro	Glu	Asp	Ile	Tyr	Gln	Met	Ala
			115				120						125		

Leu	Asn	Gln	Ala	Leu	Lys	Asp	Leu	Lys	Ala	Leu	Gly	Cys	Arg	Lys	Ala
	130					135					140				

Met	Lys	Lys	Phe	Glu	Arg	His	Thr	Leu	Leu	Glu	Tyr	Leu	Leu	Gly	Glu
	145				150					155				160	

Gly	Asn	Leu	Ser	Arg	Pro	Ala	Val	Gln	Leu	Leu	Gly	Asp	Val	Met	Ser
			165						170					175	

Glu	Asp	Gly	Phe	Phe	Tyr	Leu	Ser	Phe	Ala	Glu	Ala	Leu	Arg	Ala	Xaa
			180					185					190		

Ser	Cys	Leu	Ser	Asp	Arg	Leu	Gln	Tyr	Ser	Arg	Ile	Val	Gly	Gly	Trp
		195					200						205		

Asp	Leu	Leu	Pro	Arg	Ala	Leu	Leu	Ser	Ser	Leu	Ser	Gly	Leu	Val	Leu
	210						215					220			

Leu	Asn	Ala	Pro	Val	Val	Ala	Met	Thr	Gln	Gly	Pro	His	Asp	Val	His
	225					230				235				240	

Val	Gln	Ile	Glu	Thr	Ser	Pro	Pro	Ala	Arg	Asn	Leu	Lys	Val	Leu	Lys
				245					250					255	

Ala	Asp	Val	Val	Leu	Leu	Thr	Ala	Ser	Gly	Pro	Ala	Val	Lys	Arg	Ile
				260				265					270		

184

Thr Phe Ser
275

<210> 261
<211> 212
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (123)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 261
Leu Pro Arg His Met Gln Glu Ala Leu Arg Arg Leu His Tyr Val Pro
1 5 10 15
Ala Thr Lys Val Phe Leu Ser Phe Arg Arg Pro Phe Trp Arg Glu Glu
20 25 30
His Ile Glu Gly Gly His Ser Asn Thr Asp Arg Pro Ser Arg Met Ile
35 40 45
Phe Tyr Pro Pro Pro Arg Glu Gly Ala Leu Leu Leu Ala Ser Tyr Thr
50 55 60
Trp Ser Asp Ala Ala Ala Phe Ala Gly Leu Ser Arg Glu Glu Ala
65 70 75 80
Leu Arg Leu Ala Leu Asp Asp Val Ala Ala Leu His Gly Pro Val Val
85 90 95
Arg Gln Leu Trp Asp Gly Thr Gly Val Val Lys Arg Trp Ala Glu Asp
100 105 110
Gln His Ser Gln Gly Gly Phe Val Val Gln Xaa Pro Ala Leu Trp Gln
115 120 125
Thr Glu Lys Asp Asp Trp Thr Val Pro Tyr Gly Arg Ile Tyr Phe Ala
130 135 140
Gly Glu His Thr Ala Tyr Pro His Gly Trp Val Glu Thr Ala Val Lys
145 150 155 160
Ser Ala Leu Arg Ala Ala Ile Lys Ile Asn Ser Arg Lys Gly Pro Ala
165 170 175
Ser Asp Thr Ala Ser Pro Glu Gly His Ala Ser Asp Met Glu Gly Gln
180 185 190
Gly His Val His Gly Val Ala Ser Ser Pro Ser His Asp Leu Ala Lys
195 200 205
Glu Glu Gly Ser
210

185

<210> 262
 <211> 319
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (68)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (115)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (213)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 262
 Met Ala Pro Leu Ala Leu His Leu Leu Val Leu Val Pro Ile Leu Leu
 1 5 10 15
 Ser Leu Val Ala Ser Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro
 20 25 30
 Phe Glu Lys Cys Met Gln Asp Pro Asp Tyr Glu Gln Leu Leu Lys Val
 35 40 45
 Thr Ile Leu Glu Ala Asp Asn Arg Ile Gly Gly Arg Ile Phe Thr Tyr
 50 55 60
 Arg Asp Gln Xaa Thr Gly Trp Ile Gly Glu Leu Gly Ala Met Arg Met
 65 70 75 80
 Pro Ser Ser His Arg Ile Leu His Lys Leu Cys Gln Gly Leu Gly Leu
 85 90 95
 Asn Leu Thr Lys Phe Thr Gln Tyr Asp Lys Asn Thr Trp Thr Glu Val
 100 105 110
 His Glu Xaa Lys Leu Arg Asn Tyr Val Val Glu Lys Val Pro Glu Lys
 115 120 125
 Leu Gly Tyr Ala Leu Arg Pro Gln Glu Lys Gly His Ser Pro Glu Asp
 130 135 140
 Ile Tyr Gln Met Ala Leu Asn Gln Ala Leu Lys Asp Leu Lys Ala Leu
 145 150 155 160
 Gly Cys Arg Lys Ala Met Lys Lys Phe Glu Arg His Thr Leu Leu Glu
 165 170 175
 Tyr Leu Leu Gly Glu Gly Asn Leu Ser Arg Pro Ala Val Gln Leu Leu
 180 185 190

186

Gly Asp Val Met Ser Glu Asp Gly Phe Phe Tyr Leu Ser Phe Ala Glu
195 200 205

Ala Leu Arg Ala Xaa Ser Cys Leu Ser Asp Arg Leu Gln Tyr Ser Arg
210 215 220

Ile Val Gly Gly Trp Asp Leu Leu Pro Arg Ala Leu Leu Ser Ser Leu
225 230 235 240

Ser Gly Leu Val Leu Leu Asn Ala Pro Val Val Ala Met Thr Gln Gly
245 250 255

Pro His Asp Val His Val Gln Ile Glu Thr Ser Pro Pro Ala Arg Asn
260 265 270

Leu Lys Val Leu Lys Ala Asp Val Val Leu Leu Thr Ala Ser Gly Pro
275 280 285

Ala Val Lys Arg Ile Thr Phe Ser Pro Arg Cys Pro Ala Thr Cys Arg
290 295 300

Arg Arg Cys Gly Gly Cys Thr Thr Cys Arg Pro Pro Arg Cys Ser
305 310 315

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/15849

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) : C12N 15/11, 15/00, 15/63; C07H 21/02, 21/04 US CL : 536/23.1, 23.4; 435/320.1, 69.1 According to International Patent Classification (IPC) or to both national classification and IPC														
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 536/23.1, 23.4; 435/320.1, 69.1 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Sequence searched SEQ ID NO:11, SEQ ID NO: 103, APS, STN, CAPLUS, terms nucleic acid, express, vector, pancreas islet cell tumors.														
C. DOCUMENTS CONSIDERED TO BE RELEVANT														
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
A, P	US 5,849,498 A (BANDMAN et al) 15 December 1998, sequence listing.	1-10, 14-15, and 21												
A	US 5,670,367 A (DORNER et al) 23 September 1997, especially sequence listing.	1-10, 14-15, and 21												
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.														
<table border="0"><tr><td>* Special categories of cited documents:</td><td>*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</td></tr><tr><td>*A* document defining the general state of the art which is not considered to be of particular relevance</td><td>*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</td></tr><tr><td>*B* earlier document published on or after the international filing date</td><td>*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</td></tr><tr><td>*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</td><td>*A* document member of the same patent family</td></tr><tr><td>*O* document referring to an oral disclosure, use, exhibition or other means</td><td></td></tr><tr><td>*P* document published prior to the international filing date but later than the priority date claimed</td><td></td></tr></table>			* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family	*O* document referring to an oral disclosure, use, exhibition or other means		*P* document published prior to the international filing date but later than the priority date claimed	
* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention													
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone													
B earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art													
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family													
O document referring to an oral disclosure, use, exhibition or other means														
P document published prior to the international filing date but later than the priority date claimed														
Date of the actual completion of the international search 30 SEPTEMBER 1999		Date of mailing of the international search report 21 OCT 1999												
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer LI LEE Telephone No. (703) 308-0196												

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/15849**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-10, 14-15, and 21

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/15849

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-10, 14-15, and 21, drawn to isolated nucleic acid and expression system.

Group II, claim(s) 11-12 and 16, drawn to isolated polypeptide.

Group III, claim(s) 13, drawn to antibody.

Group IV, claim(s) 17, drawn to method for preventing a medical condition.

Group V, claim(s) 18-19, drawn to method of diagnosing a disease.

Group VI, claim(s) 20, 22, drawn to method for identifying a binding partner.

Group VII, claim 23, drawn to product produced by method of claim 20.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of Group I is considered to be isolated nucleic acid and expression system.

The special technical feature of Group II is considered to be isolated polypeptide.

The special technical feature of Group III is considered to be antibody.

The special technical feature of Group IV is considered to be a method for preventing a medical condition..

The special technical feature of Group V is considered to be method of diagnosing a disease.

The special technical feature of Group VI is considered to be a method for identifying a binding partner..

The special technical feature of Group VII is considered to be product produced by method of claim 20.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

There are 71 genes, from gene Nos 1-71.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: The 71 genes have different nucleic acid sequences and they are from different cell. Therefore they lack same special technical features.